

STIC-Biot ch/ChemLib

93 845

From: Ramirez, Delia
Sent: Wednesday, June 04, 2003 5:55 PM
To: STIC-Biotech/ChemLib
Subject: case 09/777,566

RECEIVED

JUN -5 2003

(STIC)

Hi,

I would like to request the following search (commercial and interference):

1. a standard search of seq id 1 (DNA) in the nucleic acid databases
2. a standard search of seq id 2 (Protein) in the nucleic acid databases
3. a standard search of seq id 2 (Protein) in the protein databases

Thank you,

Delia M. Ramirez, Ph.D.
Patent Examiner
Recombinant Enzymes-Art Unit 1652
USPTO
1911 S. Clark Street, Crystal Mall 1, 10D06, Mail room 10D01
Arlington, VA 22202
(703) 306-0288
delia.ramirez@uspto.gov

Edward Han
Technical Info. Specialist
STIC/Biotech
CMI 6B02 Tel: 305-9203

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 6/5/03
Date Completed: 6/13/03
Searcher Prep/Review: _____
Clerical: _____
Online time: _____

TYPE OF SEARCH:

NA Sequences: _____
AA Sequences: IFT Reverse
Structures: _____ to NA
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)

STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: Q1/Q2
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

Pending Nucleic Acid and/or Pending Amino Acid database searches now generate two sets of results. These databases were split into two parts to reduce the time needed to update the databases daily. The split freed up more machine time for processing searches.

Searches run against the Nucleic Acid Pending database produce two sets of results, with the extensions, **.rnpm** and **.rnpn**

Searches run against the Amino Acid Pending database produce two sets of results, with the extensions, **.rapm** and **.rapn**

The Pending database search results should not be left in the case because they contain data that is confidential.

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)



STIC SEARCH RESULTS

Biotech-Chem Library

Questions about the scope or the results of the search? Contact *the searcher or contact*:

Mary Hale, Information Branch Supervisor
308-4258, CM1-1E01

Voluntary Results Feedback Form

➤ I am an examiner in Workgroup: Example: 1610

➤ Relevant prior art **found**, search results used as follows:

- ☐ 102 rejection
- ☐ 103 rejection
- ☐ Cited as being of interest.
- ☐ Helped examiner better understand the invention.
- ☐ Helped examiner better understand the state of the art in their technology.

Types of relevant prior art found:

- ☐ Foreign Patent(s)
- ☐ Non-Patent Literature
(journal articles, conference proceedings, new product announcements etc.)

➤ Relevant prior art **not found**:

- ☐ Results verified the lack of relevant prior art (helped determine patentability).
- ☐ Results were not useful in determining patentability or understanding the invention.

Comments:

Drop off or send completed forms to STIC/Biotech-Chem Library CM1 - Circ. Desk



THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 06:51:00 ; Search time 64 Seconds
(Without alignments)
6339,584 Million cell updates/sec

Title: US-09-777-566a-1
Perfect score: 1323
Sequence: 1 atgaagcgcattcaatccc.....atccaccatccatccacta 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapept 1.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 682724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database:

Issued_Patents_NA: *
1: /cgn2.6/prodata/1/ina/5A.COMB.seq: *
2: /cgn2.6/prodata/1/ina/5B.COMB.seq: *
3: /cgn2.6/prodata/1/ina/6A.COMB.seq: *
4: /cgn2.6/prodata/1/ina/6B.COMB.seq: *
5: /cgn2.6/prodata/1/ina/PCUS.COMB.seq: *
6: /cgn2.6/prodata/1/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1322	99.9	1323	US-09-259-214-1	Sequence 1, Appl1
2	1322	99.9	1323	US-09-318-528-1	Sequence 1, Appl1
3	1322	99.9	1323	US-09-291-931-1	Sequence 1, Appl1
4	1210	91.5	1272	US-08-910-798-1	Sequence 1, Appl1
5	34.4	2.6	4403765	US-09-103-840A-2	Sequence 2, Appl1
6	34.4	2.6	4411529	US-09-103-840A-1	Sequence 1, Appl1
7	34	2.6	998	US-07-885-089B-5	Sequence 5, Appl1
8	33	2.5	2511	US-08-680-326-116	Sequence 116, Appl
9	33	2.5	35100	US-08-770-379-17	Sequence 17, Appl
10	33	2.5	35100	US-08-757-669A-17	Sequence 17, Appl
11	33	2.5	35100	US-09-230-371A-17	Sequence 17, Appl
12	31.8	2.4	4403765	US-09-103-840A-2	Sequence 2, Appl1
13	31.8	2.4	4411529	US-09-103-840A-1	Sequence 1, Appl1
14	31.6	2.4	8460	US-08-469-005A-9	Sequence 9, Appl1
15	31.6	2.4	8460	US-09-261-907-1	Sequence 1, Appl1
16	31.2	2.4	1034	US-09-072-596-312	Sequence 312, App
17	31	2.3	3147	US-08-441-430-3	Sequence 3, Appl1
18	31	2.3	4488	US-08-441-430-1	Sequence 1, Appl1
19	30.4	2.3	418	US-09-072-596-302	Sequence 302, App
20	30.4	2.3	8252	US-08-046-585-15	Sequence 15, Appl
21	30.4	2.3	8252	US-08-393-703-15	Sequence 15, Appl
22	30.4	2.3	8252	PCT-US93-11721-15	Sequence 15, Appl
23	29.8	2.3	9711	US-08-961-527-167	Sequence 167, App
24	29.6	2.2	150	US-08-659-473-8	Sequence 8, Appl1
25	29.6	2.2	1005	US-08-875-062-4	Sequence 4, Appl1
26	29.6	2.2	1521	US-08-496-855A-3	Sequence 3, Appl1
27	29.6	2.2	1521	US-07-938-154-9	Sequence 9, Appl1

C 28	29.6	2.2	1521	5	PCT-US91-02311-9	Sequence 9, Appl1
C 29	29.6	2.2	2448	4	US-08-487-596-13	Sequence 13, Appl
C 30	29.6	2.2	2450	2	US-08-466-589-9	Sequence 9, Appl1
C 31	29.6	2.2	2450	2	US-08-700-636-8	Sequence 9, Appl1
C 32	29.6	2.2	2450	3	US-08-467-574-9	Sequence 9, Appl1
C 33	29.6	2.2	2450	3	US-09-217-345-9	Sequence 9, Appl1
C 34	29.6	2.2	6379	1	US-08-499-215-1	Sequence 1, Appl1
C 35	29.2	2.2	508	3	US-08-669-408B-3	Sequence 3, Appl1
C 36	29.2	2.2	1555	3	US-08-669-408B-9	Sequence 9, Appl1
C 37	29	2.2	1548	2	US-08-762-106-5	Sequence 5, Appl1
C 38	29	2.2	1548	4	US-09-320-774-5	Sequence 5, Appl1
C 39	29	2.2	1581	2	US-08-762-106-6	Sequence 6, Appl1
C 40	29	2.2	1581	4	US-09-320-774-6	Sequence 6, Appl1
C 41	29	2.2	3531	1	US-08-629-600-1	Sequence 1, Appl1
C 42	29	2.2	3531	4	US-09-076-761-1	Sequence 1, Appl1
C 43	29	2.2	7680	4	US-09-210-748A-3	Sequence 3, Appl1
C 44	29	2.2	56516	2	US-08-996-306-1	Sequence 1, Appl1
C 45	29	2.2	56516	4	US-09-338-907-1	Sequence 1, Appl1

ALIGNMENTS

91777-566

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 6110719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHRYASE
FILE REFERENCE: DIVER1370-1
CURRENT APPLICATION NUMBER: US/09/259, 214A
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: 08/910, 798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1320)
NAME/KEY: misc-feature
LOCATION: (1)..(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1

Query Match 99.9%; Score 1322; DB 3; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGAACGAGATCTTAATCCATTTTATCTTCTTGATTCGTTAACCCGCAATGCA	60
DB	1	ATGAACGAGATCTTAATCCATTTTATCTTCTTGATTCGTTAACCCGCAATGCA	60
QY	61	TTGCGTCAAGTGAAGCGGAGAGCTGGAAGTGTGATTTGCTCACTGCTATGCT	120
DB	61	TTGCGTCAAGTGAAGCGGAGAGCTGGAAGTGTGATTTGCTCACTGCTATGCT	120
QY	121	GTGCGTCTCCAAACCAAGGCGCAGCACTATGAGATGTCACCCAGAGCAATGGCCA	180
DB	121	GTGCGTCTCCAAACCAAGGCGCAGCACTATGAGATGTCACCCAGAGCAATGGCCA	180
QY	181	ACCTGCGGCTAAACTGGTGTGCTGACACCGGCGGTGAGTAAATCGCTATCTC	240
DB	181	ACCTGCGGCTAAACTGGTGTGCTGACACCGGCGGTGAGTAAATCGCTATCTC	240
QY	241	GGACATTACCAACCCAGCGCTGTGTAGCCGAGCATTTGCTGGCGAAAAGGCTGCCG	300
DB	241	GGACATTACCAACCCAGCGCTGTGTAGCCGAGCATTTGCTGGCGAAAAGGCTGCCG	300

301 CAGTCTGCTGAGTGGCGGATTAATGCTGATGTCGACGAGCGTACCCGTAACAGCGCA 360
 301 CAGTCTGCTGAGTGGCGGATTAATGCTGATGTCGACGAGCGTACCCGTAACAGCGCA 360
 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATACCGTACATACCCAGGCGATACG 420
 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATACCGTACATACCCAGGCGATACG 420
 421 TCCAGTCCCGGATCGGTTATTAATCCTTAATAAACTGGCGTTGGCAACTGGATTAACGCG 480
 421 TCCAGTCCCGGATCGGTTATTAATCCTTAATAAACTGGCGTTGGCAACTGGATTAACGCG 480
 421 TCCAGTCCCGGATCGGTTATTAATCCTTAATAAACTGGCGTTGGCAACTGGATTAACGCG 480
 481 AACGTGACTGACGCGGATCTCAGCAGGCGAGGCGTCAATGCTGACTTACCGGCGAT 540
 481 AACGTGACTGACGCGGATCTCAGCAGGCGAGGCGTCAATGCTGACTTACCGGCGAT 540
 541 CGGCAAAAGCGGTTTCGGGAACGAGGCGTCTAATTTCCGCAATCAAACTGTGCG 600
 541 CGGCAAAAGCGGTTTCGGGAACGAGGCGTCTAATTTCCGCAATCAAACTGTGCG 600
 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGCAATACCGGCACTC 660
 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGCAATACCGGCACTC 660
 661 AAGGTGAGCGCGGACATGCTCATTAACCGGCTGGGTAAGCTTCGCAATGCTGACG 720
 661 AAGGTGAGCGCGGACATGCTCATTAACCGGCTGGGTAAGCTTCGCAATGCTGACG 720
 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCGGCGGAGGTAAGTCAACG 780
 721 GAGATATTTCTCTGCAACAGCAGAGGAAATGCGGAGCGGCGGAGGTAAGTCAACG 780
 781 GATTACACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACAA 840
 781 GATTACACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACAA 840
 841 GATTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACAA 840
 841 GATTCACACAGTGGAAACACCTTGCTAAGTTGCTAATACGCGCAATTTATTTGCTACAA 840
 841 CGCAGCGCAGAGTGGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 841 CGCAGCGCAGAGTGGCGCGGACGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 900
 901 TTGAGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960
 901 TTGAGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960
 901 TTGAGCGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 960
 961 TTTAATGCGGAGACAGATATTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGGCGG 1020
 961 TTTAATGCGGAGACAGATATTAATCTGGCAAACTCTGGCGGCGGCGGCGGCGGCGGCGG 1020
 1021 ACAGTCCCGGCTACGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 1021 ACAGTCCCGGCTACGCGGATTAACAGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
 1081 CGTGGCGCTAACGAGTAACAGCAGTGTGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 1081 CGTGGCGCTAACGAGTAACAGCAGTGTGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 1081 CGTGGCGCTAACGAGTAACAGCAGTGTGAGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 1141 CAGATGCGCTGATTAACAGCGCGCTGCTGATTAATACGCGCGCGGAGAGAGTGAACGTAAC 1200
 1141 CAGATGCGCTGATTAACAGCGCGCTGCTGATTAATACGCGCGCGGAGAGAGTGAACGTAAC 1200
 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGGCGATGTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGAGGCGATGTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 1261 ATCGTGAATGAAGACGATACGCGCGTGTGAGTTGAGATCTCAATCAACATACCATAC 1320
 1261 ATCGTGAATGAAGACGATACGCGCGTGTGAGTTGAGATCTCAATCAACATACCATAC 1320
 1321 TAA 1323
 1321 TAA 1323
 1321 TAA 1323

RESULT 2
 US-09-318-528-1
 : Sequence 1, Application US/09318528
 : Patent No. 6183740
 : GENERAL INFORMATION:
 : APPLICANT: Kretz, Keith
 : TITLE OF INVENTION: NOVEL PHYTASE
 : FILE REFERENCE: 09010/029003
 : CURRENT APPLICATION NUMBER: US/09/318,528
 : EARLIER FILING DATE: 1999-05-25
 : EARLIER APPLICATION NUMBER: 09/291,931
 : EARLIER FILING DATE: 1999-04-13
 : EARLIER APPLICATION NUMBER: 08/910,798
 : EARLIER FILING DATE: 1997-08-13
 : EARLIER APPLICATION NUMBER: 09/259,214
 : EARLIER FILING DATE: 1999-03-01
 : NUMBER OF SEQ ID NOS: 4
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO 1
 : LENGTH: 1323
 : TYPE: DNA
 : ORGANISM: Escherichia coli
 : FEATURE:
 : NAME/KEY: CDS
 : LOCATION: (1)...(1320)
 : NAME/KEY: misc_feature
 : LOCATION: (1)...(1323)
 : OTHER INFORMATION: n = A,T,C or G
 : US-09-318-528-1

Query Match 99.9%; Score 1322; DB 4; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGGAGATCTTAATCCCAATTTATCTCTGATTCGTTAACCCGCAATCTCA 60
 1 ATGAAGGAGATCTTAATCCCAATTTATCTCTGATTCGTTAACCCGCAATCTCA 60
 61 TTCGCTCAGAGTGAACCGGAGCTGAAGCTGAAGTGTGATGTCAGTCTCATG 120
 61 TTCGCTCAGAGTGAACCGGAGCTGAAGCTGAAGTGTGATGTCAGTCTCATG 120
 121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
 121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
 121 GTGCGTGTCCCAACCAAGCCAGCACTGATGCAAGATGTCACCCGAGAGCATGGCA 180
 181 ACCTGGCGGTAATAAAGTGGTGGCTGACACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 181 ACCTGGCGGTAATAAAGTGGTGGCTGACACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 181 ACCTGGCGGTAATAAAGTGGTGGCTGACACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240
 241 GGACATTAACCAAGCGCAGCGCTGTGTAGCCGACGAGATTGCTGGCGAAAGGCGTCCCG 300
 241 GGACATTAACCAAGCGCAGCGCTGTGTAGCCGACGAGATTGCTGGCGAAAGGCGTCCCG 300
 301 CAGTCTGCTCAGTGTGCGCATTTATGCTGATGTGCGACGAGCGTACCCGTAACAGCGCAA 360
 301 CAGTCTGCTCAGTGTGCGCATTTATGCTGATGTGCGACGAGCGTACCCGTAACAGCGCAA 360
 301 CAGTCTGCTCAGTGTGCGCATTTATGCTGATGTGCGACGAGCGTACCCGTAACAGCGCAA 360
 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTAACCCAGGCGATACG 420
 361 GCCTTCGCGCGCGGCTGGACCTGACTGTGCAATTAACCGTAACCCAGGCGATACG 420
 421 TCCAGTCCCGGATCGGTTATTAATCCTTAATAAACTGGCGTTGGCAACTGGATTAACGCG 480
 421 TCCAGTCCCGGATCGGTTATTAATCCTTAATAAACTGGCGTTGGCAACTGGATTAACGCG 480
 481 AACGTGACTGACGCGGATCTCAGCAGGCGAGGCGTCAATGCTGACTTACCGGCGAT 540
 481 AACGTGACTGACGCGGATCTCAGCAGGCGAGGCGTCAATGCTGACTTACCGGCGAT 540
 541 CGGCAAAAGCGGTTTCGGGAACGAGGCGTCTAATTTCCGCAATCAAACTGTGCG 600
 541 CGGCAAAAGCGGTTTCGGGAACGAGGCGTCTAATTTCCGCAATCAAACTGTGCG 600

```

OY 601 CTTAAACGTGAGAAACAGGACGAAAGCTTGTTCATTAACGACGATTAACATCGGAATC 660
    |||||||
DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTTGTTCATTAACGACGATTAACATCGGAATC 660
OY 661 AAGGTGAGGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGTGTAG 720
    |||||||
DB 661 AAGGTGAGGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGTGTAG 720
OY 721 GAGATATTTCTCTCGACACAGACAGGAATCCGGAGCCGGGGTGGGAAGATACCC 780
    |||||||
DB 721 GAGATATTTCTCTCGACACAGACAGGAATCCGGAGCCGGGGTGGGAAGATACCC 780
OY 781 GATTACACCACTGGAAACCTTGTGTAAGTTGCATTAACGCGCAATTTATTTGCTACAA 840
    |||||||
DB 781 GATTACACCACTGGAAACCTTGTGTAAGTTGCATTAACGCGCAATTTATTTGCTACAA 840
OY 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCGTTATTTGATTTGATCAATGCGAGCG 900
    |||||||
DB 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCGTTATTTGATTTGATCAATGCGAGCG 900
OY 901 TTGAGAGCCCATCCACCGGAAAAAGCGGTATGCTGTACATTAACCACTTCAGTACG 960
    |||||||
DB 901 TTGAGAGCCCATCCACCGGAAAAAGCGGTATGCTGTACATTAACCACTTCAGTACG 960
OY 961 TTTATTTCCGCGACAGATTAATCTGGAATTCGCGGCGGCGACCTGGAGCTCAACTGG 1020
    |||||||
DB 961 TTTATTTCCGCGACAGATTAATCTGGAATTCGCGGCGGCGACCTGGAGCTCAACTGG 1020
OY 1021 ACGCTTCCGCGTACGCGGATTAACACGCGCCAGGTGTGTAACCTGTTTGAACCGCTG 1080
    |||||||
DB 1021 ACGCTTCCGCGTACGCGGATTAACACGCGCCAGGTGTGTAACCTGTTTGAACCGCTG 1080
OY 1081 CCGTGGCTTAAGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACG 1140
    |||||||
DB 1081 CCGTGGCTTAAGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACG 1140
OY 1141 CAGATGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGG 1200
    |||||||
DB 1141 CAGATGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGG 1200
OY 1201 CTGGCAGAGTGTGAAGAGCGAATGCGGAGCATGTTCGTTGCGAGGTTTACGCAA 1260
    |||||||
DB 1201 CTGGCAGAGTGTGAAGAGCGAATGCGGAGCATGTTCGTTGCGAGGTTTACGCAA 1260
OY 1261 ATCGTAATGAAGACGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACG 1320
    |||||||
DB 1261 ATCGTAATGAAGACGATTAACGCGGATTAACGCGGATTAACGCGGATTAACGCGGATTAACG 1320
OY 1321 TAA 1323
    |||
DB 1321 TAA 1323

```

```

RESULT 3
US-09-291-931-1
: Sequence 1, Application us/09291931A
: Patent No. 6190897
: GENERAL INFORMATION:
: APPLICANT: Kretz, Keith
: TITLE OF INVENTION: NOVEL PHYTASE
: FILE REFERENCE: 09010/029003
: CURRENT APPLICATION NUMBER: US/09/291, 931A
: CURRENT FILING DATE: 1999-04-13
: EARLIER APPLICATION NUMBER: 08/910, 798
: EARLIER FILING DATE: 1997-08-13
: EARLIER APPLICATION NUMBER: 09/259, 214
: EARLIER FILING DATE: 1999-03-01
: NUMBER OF SEQ ID NOS: 4
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 1
: LENGTH: 1323
: TYPE: DNA

```

```

: ORGANISM: Escherichia coli
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (1)...(1320)
: NAME/KEY: misc_feature
: LOCATION: (1)...(1323)
: OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

```

```

Query Match      99.9%; Score 1322; DB 4; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 ATGAAGCATCTTATCCATTTTATCTCTGATTCGCTTAACCCCGCAATGTGA 60
    |||||||
DB 1 ATGAAGCATCTTATCCATTTTATCTCTGATTCGCTTAACCCCGCAATGTGA 60
OY 61 TTGCTCAGAGTGAAGCGGAGGTGAAGTGTGTGATTTGTCAGTGTATGTGT 120
    |||||||
DB 61 TTGCTCAGAGTGAAGCGGAGGTGAAGTGTGTGATTTGTCAGTGTATGTGT 120
OY 121 GTGCGTCTCAACCAAGGCGACGCACTGATGACAGATGTACCCCAAGCATGCGCA 180
    |||||||
DB 121 GTGCGTCTCAACCAAGGCGACGCACTGATGACAGATGTACCCCAAGCATGCGCA 180
OY 181 ACCTGCGGTAAACCTGGGTTGGTGAACCGCGGAGTGTGATGCTAAATCGCTATCTC 240
    |||||||
DB 181 ACCTGCGGTAAACCTGGGTTGGTGAACCGCGGAGTGTGATGCTAAATCGCTATCTC 240
OY 241 GGCATTTTACCAACGCGGCGTGTGATGCGGAGATTTCTGCGGAAAAAGGCTGCGCG 300
    |||||||
DB 241 GGCATTTTACCAACGCGGCGTGTGATGCGGAGATTTCTGCGGAAAAAGGCTGCGCG 300
OY 301 CAGTGTGTCAAGTGTGCGGATTTATGCTGATGTGACAGAGCTTACCGGTAAACAGCGCA 360
    |||||||
DB 301 CAGTGTGTCAAGTGTGCGGATTTATGCTGATGTGACAGAGCTTACCGGTAAACAGCGCA 360
OY 361 GCGTTGCGCGGCGGCGTGCACCTGACCTGACATTAACGATTAACGATTAACGATTAACG 420
    |||||||
DB 361 GCGTTGCGCGGCGGCGTGCACCTGACCTGACATTAACGATTAACGATTAACGATTAACG 420
OY 421 TCCAGTCCGCGATCCGTTATTTATTCCTCTTAAATAATGCGGTTGCCAATGATAACGCG 480
    |||||||
DB 421 TCCAGTCCGCGATCCGTTATTTATTCCTCTTAAATAATGCGGTTGCCAATGATAACGCG 480
OY 481 AACGTACTGACGCGATCTCTACAGAGGAGGATCAATGTCTACTTTACCGGGCAT 540
    |||||||
DB 481 AACGTACTGACGCGATCTCTACAGAGGAGGATCAATGTCTACTTTACCGGGCAT 540
OY 541 CGGCAAAAGGCGTTTGGCGAATGGAAGGCGTCTTAATTTTCCGCAATCAACTGTGCG 600
    |||||||
DB 541 CGGCAAAAGGCGTTTGGCGAATGGAAGGCGTCTTAATTTTCCGCAATCAACTGTGCG 600
OY 601 CTTAAACGTGAGAAACAGGACGAAAGCTTGTTCATTAACGACGATTAACATCGGAATC 660
    |||||||
DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTTGTTCATTAACGACGATTAACATCGGAATC 660
OY 661 AAGGTGAGGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGTGTAG 720
    |||||||
DB 661 AAGGTGAGGCGCCGACAAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGTGTAG 720
OY 721 GAGATATTTCTCTCGACACAGACAGGAATCCGGAGCCGGGGTGGGAAGATACCC 780
    |||||||
DB 721 GAGATATTTCTCTCGACACAGACAGGAATCCGGAGCCGGGGTGGGAAGATACCC 780
OY 781 GATTACACCACTGGAAACCTTGTGTAAGTTGCATTAACGCGCAATTTATTTGCTACAA 840
    |||||||
DB 781 GATTACACCACTGGAAACCTTGTGTAAGTTGCATTAACGCGCAATTTATTTGCTACAA 840
OY 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCGTTATTTGATTTGATCAATGCGAGCG 900
    |||||||
DB 841 CGCAGCGCAGAGGTTGCCCGCAGCCGCGCACCCGTTATTTGATTTGATCAATGCGAGCG 900

```

```

QY 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 960
Db 901 TTTGACGCCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 960
QY 961 TTTATTCGCGGACAGATTAATCTGGCAAAATCTGGCGCGGACGACTGAGCACTCAACG 1020
Db 961 TTTATTCGCGGACAGATTAATCTGGCAAAATCTGGCGCGGACGACTGAGCACTCAACG 1020
QY 1021 ACAGTTCGCGGACGCGGATTAACAGCGCGGACGAGTGTGAACGCTGTTGAACGCTG 1080
Db 1021 ACAGTTCGCGGACGCGGATTAACAGCGCGGACGAGTGTGAACGCTGTTGAACGCTG 1080
QY 1081 CGTGGCGTAAAGCGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1140
Db 1081 CGTGGCGTAAAGCGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1140
QY 1141 CAGATGCGGTAAAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1200
Db 1141 CAGATGCGGTAAAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1200
QY 1201 CTGGCAGAGTGTAAAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1260
Db 1201 CTGGCAGAGTGTAAAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1260
QY 1261 ATCGTGAATGAAGCAGCATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1320
Db 1261 ATCGTGAATGAAGCAGCATTAACAGCGCGGATTAACAGCGCGGATTAACAGCGCGGATTAAC 1320
QY 1321 TAA 1323
Db 1321 TAA 1323

```

```

RESULT 4
US-08-910-798-1
: Sequence 1, Application us/08910798
: Patent No. 5876997
: GENERAL INFORMATION:
: APPLICANT: KRETEZ
: TITLE OF INVENTION: NOVEL PHYTASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: IBM PC compatible
: SOFTWARE: Patent Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: FILING DATE: August 13, 1997
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HAILE, PH.D., LISA A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09010/029001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5070
: TELEFAX: 619/678-5070
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1272 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: PHYTASE

```

```

: FEATURE:
: NAME/KEY:
: LOCATION:
: US-08-910-798-1
Query Match 91.5%, Score 1210, DB 2, Length 1272:
Best Local Similarity 96.1%, Pred. No. 0;
Matches 1272; Conservative 0; Mismatches 0; Indels 51; Gaps 1;
QY 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
Db 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGCTTAACCCGCAATCTGCA 60
QY 61 TTCGCTCAGAGTGAAGCGGACCTGAAGCTGGAAGTGTGATTTGTCAGTCTATGCT 120
Db 61 TTCGCTCAGAGTGAAGCGGACCTGAAGCTGGAAGTGTGATTTGTCAGTCTATGCT 120
QY 121 GTGCGTGTCCCAACCAAGGCGCAGCACTGATGAGATGTCACCCGACAGCATGGCCA 180
Db 121 GTGCGTGTCCCAACCAAGGCGCAGCACTGATGAGATGTCACCCGACAGCATGGCCA 180
QY 181 ACCTGGCGGTTAAAGTGGCTGTGACACCGCGGCTGAGCTAATGCGCTATCTC 240
Db 181 ACCTGGCGGTTAAAGTGGCTGTGACACCGCGGCTGAGCTAATGCGCTATCTC 240
QY 241 GGACATTACCAACGCGCAGCGCTGTGTAGCCGAGCGGATTTGCGGCAAAAGGCTGCCG 300
Db 241 GGACATTACCAACGCGCAGCGCTGTGTAGCCGAGCGGATTTGCGGCAAAAGGCTGCCG 300
QY 301 CAGTGTGTGACGTCGCGGATTAATGCTGATGTCACGAGCGTACCCTTAACAGCGGAA 360
Db 301 CAGTGTGTGACGTCGCGGATTAATGCTGATGTCACGAGCGTACCCTTAACAGCGGAA 360
QY 361 GCTTGGCGCGCGGCTGCGACCTGACTGTGCAATTAACCGTATACCGACGAGATACG 420
Db 361 GCTTGGCGCGCGGCTGCGACCTGACTGTGCAATTAACCGTATACCGACGAGATACG 420
QY 421 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGCG 480
Db 421 TCAGTCCCGATCGGTTATTTAATCTCTAATAAACTGGCTTCCCACTGATTAACGCG 480
QY 481 AACGTGACGAGCGGATCTCAGAGAGGAGGAGGATTAATGCTGATTTACGGGCGAT 540
Db 481 AACGTGACGAGCGGATCTCAGAGAGGAGGAGGATTAATGCTGATTTACGGGCGAT 540
QY 541 CGGCAAAAGCGGCTTCCGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
Db 541 CGGCAAAAGCGGCTTCCGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
QY 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACGATGGAATC 660
Db 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACGATGGAATC 660
QY 661 AAGGTGAGCGCGGCAATGCTCTAATTAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 720
Db 661 AAGGTGAGCGCGGCAATGCTCTAATTAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 720
QY 721 GAGATATTTTCTCCGCAACAGCAGGAGGAAATGCGGAGCGGAGTGGGGAAGATACG 780
Db 721 GAGATATTTTCTCCGCAACAGCAGGAGGAAATGCGGAGCGGAGTGGGGAAGATACG 780
QY 781 GATTACACAGGAGGAAACACCTTGTAGTTTGCATTAACGCGCAATTTTATTGCTACAA 840
Db 781 GATTACACAGGAGGAAACACCTTGTAGTTTGCATTAACGCGCAATTTTATTGCTACAA 840
QY 841 CGCAGCGCAGAGGTTGCGCGGAGCGGAGCGGCAACCCGTTATTTGATTTGATGAGCGG 900
Db 841 CGCAGCGCAGAGGTTGCGCGGAGCGGAGCGGCAACCCGTTATTTGATTTGATGAGCGG 900
QY 901 TTTGACGCCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 960
Db 901 TTTGACGCCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACTG 960

```

```

Db      3834515  CCGGGGCTGGCGGCTTCCCGGGTCCCGACGGGTGCGTACCGCCGCTGTGG 3834567

RESULT 6
US-09-103-840A-1
: Sequence 1, Application US/09103840A
: Patent No. 6294328
:
GENERAL INFORMATION:
: APPLICANT: FLEISCHMAN, Robert D.
: APPLICANT: WHITE, Owen R.
: APPLICANT: FRASER, Claire M.
: APPLICANT: VENTER, John C.
: TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
: FILE REFERENCE: 24366-2007.00
: CURRENT APPLICATION NUMBER: US/09/103.840A
: CURRENT FILING DATE: 1998-06-24
: NUMBER OF SEQ ID NOS: 2
: SOFTWARE: patentln Ver. 2.1
: SEQ ID NO 1
:
LENGTH: 4411529
: TYPE: DNA
: ORGANISM: Mycobacterium tuberculosis
: OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

QY 148 CTGATGCAAGATGTCTACCCCGACAGCAGTGGCCATGCGCCGATGAAAACCTGGGTTGGCTG 207
 Db 3848436 CTGGCGGGGGGATCCACCGGTGCGCGCTCACGACGACGCCGATTCATCTGGACTC 3848495
 QY 208 ACACCGGGNGGTGTGAGCTAATGCGCTATCTCGACATTACCAACGCCACGCTGAGTA 267
 Db 3848496 GCTGCACAGGTAGATGCCCTCTGTGAGCAGCGTGGCATATACCAACGCCCGGCGCATGTC 3848555
 QY 268 GCCGACGAGATTGC-TGGCGAAAAAGGGCTGCCCGCAGTCTGGTCAGGTGCGGATTATATGC 326
 Db 3848556 AGCGAAGCGCTGCGCAGTTCCACACCGCGCAGTACTGCTCAAGTTCTGGCGCGACTCACGCG 3848615
 QY 327 TGAGTGCAGCAGCGCTACCCGTAATAACAGCGGAAGCCTTGGCGCCCGGCGCTGG 379
 Db 3848616 CCGCGCGCTGCGCGCTCTCCGCGGCTCCGACGAGGTGGTGGTACCGCCGCTGGTGG 3848668
 RESULT 7
 US-07-885-089B-5/c
 : Sequence 5, Application US/07885089B
 : Patent No. 5830995
 : GENERAL INFORMATION:
 : APPLICANT: Shoyab, Mohammed
 : APPLICANT: McDonald, Vicki L.

APPLICANT: Plowman, Gregory D.
TITLE OF INVENTION: AMPHIREGULINS: A FAMILY OF
TITLE OF INVENTION: HEPARIN-BINDING EPITHELIAL CELL GROWTH FACTORS
NUMBER OF SEQUENCES: 37
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/885,089B

Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGGCAACTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGAGAAATGTAACCTCCGTGGCAGCAGCCCAAGAGCAGGAAATA 12975

Qy 519 AATTCCTGACTTTACCGGGCAGCAACGGCGTTTGGCCGAACGGAAGGGTCTTAA 578
12974 CCCTGATCTGTTGGCCGTCCTGACCTACCGGAGGAGTCTTGGCCAGCTTGGCTA 12915

Db 579 TTTTCCGCAATCAACTTGTGCTTAAACGTGAGAGAACGAGAAAGC 627
12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAGCAGAAAGATCATAAC 12866

RESULT 10
US-08-757-669A-17/c
Sequence 17, Application US/08757669A
Patent No. 6183751

GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
TITLE OF INVENTION: SEQUENCES AND USES THEREOF
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/757,669A
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 45185-F
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 35100 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-757-669A-17

Query Match 2.5%; Score 33; DB 4; Length 35100;
Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGGCAACTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGAGAAATGTAACCTCCGTGGCAGCAGCCCTCAGAGCAGGAAATA 12975

Qy 519 AATTCCTGACTTTACCGGGCAGCAACGGCGTTTGGCCGAACGGAAGGGTCTTAA 578
12914 CCCTGATCTGTTGGCCGTCCTGACCTACCGGAGGAGTCTTGGCCAGCTTGGCTA 12915

Qy 579 TTTTCCGCAATCAACTTGTGCTTAAACGTGAGAGAACGAGAAAGC 627
12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAGCAGAAAGATCATAAC 12866

RESULT 11
US-09-230-371A-17/c
Sequence 17, Application US/09230371A
Patent No. 6348586

GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
TITLE OF INVENTION: USES THEREOF
FILE REFERENCE: 45185-G-PCT-US
CURRENT APPLICATION NUMBER: US/09/230,371A
CURRENT FILING DATE: 1999-11-17
PRIOR APPLICATION NUMBER: PCT/US97/13346
PRIOR FILING DATE: 1997-07-22
NUMBER OF SEQ ID NOS: 30
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO: 17
LENGTH: 35100
TYPE: DNA
ORGANISM: Kaposi's sarcoma-associated herpesvirus
US-09-230-371A-17

Query Match 2.5%; Score 33; DB 4; Length 35100;
Best Local Similarity 49.7%; Pred. No. 6.9;
Matches 84; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

Db 459 CGTTGGCAACTGATTAACGGCAACGCTGACGCGATCCTCAGCAGGCGAGAGGCTC 518
13034 CTTTGGGACGGGAGAAATGTAACCTCCGTGGCAGCAGCCCTCAGAGCAGGAAATA 12975

Qy 519 AATTCCTGACTTTACCGGGCAGCAACGGCGTTTGGCCGAACGGAAGGGTCTTAA 578
12974 CCCTGATCTGTTGGCCGTCCTGACCTACCGGAGGAGTCTTGGCCAGCTTGGCTA 12915

Db 579 TTTTCCGCAATCAACTTGTGCTTAAACGTGAGAGAACGAGAAAGC 627
12914 TTTCGAGATCTCAACATGGGTCTGAACCGTAGCAGAAAGATCATAAC 12866

RESULT 12
US-09-103-840A-2/c
Sequence 2, Application US/09103840A
Patent No. 6294328

GENERAL INFORMATION:
APPLICANT: FLEISCHMAN, Robert D.
APPLICANT: WHITE, Owen R.
APPLICANT: FRASER, Claire M.
APPLICANT: VENTER, John C.
TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
TITLE OF INVENTION: TUBERCULOSIS
FILE REFERENCE: 24366-2007.00
CURRENT APPLICATION NUMBER: US/09/103,840A
CURRENT FILING DATE: 1998-06-24
NUMBER OF SEQ ID NOS: 2
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO: 2
LENGTH: 4403765
TYPE: DNA
ORGANISM: Mycobacterium tuberculosis
FEATURE:
OTHER INFORMATION: CDC 1551
OTHER INFORMATION: "n" bases at various positions throughout the sequence
OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

```

Query Match      2.4%; Score 31.8; DB 4; Length 4403765;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 93; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 196 CTGGCTTGGCTGACACCGGNGTGTGAGCTAATGCGCCATCTCGACATTACCAACGC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926394 CTGGCGGGGTGTCAGCGGGTGTGCTGCTGCGCGGGGCTGGCGCCCTCCAAACCAACAA 926335
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 CAGCGCTCTGTAGCCACGAGATTTGTGCGGAAAAAGGCGTCCCGCATCTGTCAGATC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926334 CAGCCACCGAGTCCCGCCGACGCGGGTGTGCGCGGGGGGATTCGCCATTCGGCCACTGCC 926275
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 GCGATTATTGCTGATTCGACGACGCTACCCGTAACACGCGAAGCCTTCCGCCCGGG 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926274 GCGGTTCGCGCGCTTGCCTGATCAGCCGCGCCCGCGCGCGCGCGCGCGCGCGCTT 926215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 CTGGCAGCTGACGTGTG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926214 CCGCGCCCGCCCGCGGTG 926199
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 13

```

US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6293328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

```

```

Query Match      2.4%; Score 31.8; DB 4; Length 4411529;
Best Local Similarity 47.4%; Pred. No. 67;
Matches 93; Conservative 0; Mismatches 103; Indels 0; Gaps 0;

QY 196 CTGGCTTGGCTGACACCGGNGTGTGAGCTAATGCGCCATCTCGACATTACCAACGC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926495 CTGGCGGGGTGTCAGCGGGTGTGCTGCTGCGCGGGGCTTGGCGCCCTCCAAACCAACAA 926436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 256 CAGCGCTCTGTAGCCACGAGATTTGTGCGGAAAAAGGCGTCCCGCATCTGTCAGATC 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926435 CAGCCACCGAGTCCCGCCGACGCGGGTGTGCGCGGGGGGATTCGCCATTCGGCCACTGCC 926376
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 316 GCGATTATTGCTGATTCGACGACGCTACCCGTAACACGCGAAGCCTTCCGCCCGGG 375
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926375 GCGGTTCGCGCGCTTGCCTGATCAGCCCGCGCGCGCGCGCGCGCGCGCGCGCTT 926316
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 376 CTGGCAGCTGACGTGTG 391
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 926315 CCGCGCCCGCCCGCGGTG 926300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14

```

US-08-469-005A-9/C
; Sequence 9, Application US/08469005A
; Patent No. 565874
; GENERAL INFORMATION:
; APPLICANT: KUHAIJDA, FRANCIS P.
; APPLICANT: PASTERNAK, GARY A.
; TITLE OF INVENTION: CANCER RELATED ANTIGEN

```

```

; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BAKER & BOTTS, L.L.P.
; STREET: 1299 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004-2400
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,005A
; FILING DATE: 05-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/188,426
; FILING DATE: 24-JAN-1994
; APPLICATION NUMBER: 08/096,908
; FILING DATE: 26-JUL-1993
; APPLICATION NUMBER: 07/917,716
; FILING DATE: 24-JUL-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Posorske, Laurence H
; REGISTRATION NUMBER: 34,698
; REFERENCE/DOCKET NUMBER: 062482-0113
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-639-7700
; TELEFAX: 202-639-7890
; TELEX:
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8460 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE:
; ORIGINAL SOURCE:
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 124...7650
; OTHER INFORMATION:
US-08-469-005A-9

```

```

Query Match      2.4%; Score 31.6; DB 1; Length 8460;
Best Local Similarity 47.9%; Pred. No. 9.1;
Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

QY 712 ATGCTGACGAGATTTCTCTGCACACAGGAGATCGCGAGCGGGGTGGGA 771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7604 ATGCTGATGATGACACTCAGACCGCTGCTCCACAGCAGCTGGGTGATACCTCGATG 7545
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 772 AGGATCACCGATTCACACAGAGTGGAACACCTTCTTAAGTTGATTAACGCGCAATTTAT 831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7544 ACGTGACGAGATTTCTCCGTCGCAATCTGGGAGAGTGTGAATCCGCGCAGGTCC 7485
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 832 TTGCTACACGACGACGAGGTTGCCCGGAGCGCCGCCACCCCGTTATGAGATTGATC 891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7484 TCGCCGTAGCGCACCCGCTTGGCGCGAGTAGTACAGCTTGCCATGTACTTGGCC 7425
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 892 ATGACAGCT 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 7424 TTGGGTGTGT 7415
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 15

```

US-09-261-907-1/C
; Sequence 1, Application US/09261907A

```

; Patent No. 6294364
; GENERAL INFORMATION:
; APPLICANT: ELLIS, CATHERINE
; APPLICANT: LONSDALE, JOHN
; APPLICANT: BERGMA, DEK J.
; APPLICANT: MOONEY, JEFFREY L.
; APPLICANT: DEPIERA, MEGAN E.
; APPLICANT: CHAPMAN, CONRAD
; TITLE OF INVENTION: HUMAN FAS
; FILE REFERENCE: GP-70603
; CURRENT APPLICATION NUMBER: US/09/261,907A
; CURRENT FILING DATE: 1999-03-03
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 8519
; TYPE: DNA
; ORGANISM: HOMO SAPIENS
; US-09-261-907-1

Query Match 2.4%; Score 31.6; DB 4; Length 8519;

Best Local Similarity 47.9%; Pred.No.9.1;

Matches 91; Conservative 0; Mismatches 99; Indels 0; Gaps 0;

OY 712 ATGCTGACGAGATATTTCTCTGCACACAGCAGAGGAATGCCGAGCCGGGGTGGGA 771
|||||
Db 7642 ATGCTGATGATGACTCCAGCGCGCTGCCCTCCAGCAGCGGTGGGTGTCACCTCGATG 7583
OY 772 AGGATCACCAGATTACACACAGTGAACACCTTGCTAGTTGCATTAACGCGCAATTTAT 831
|||
Db 7582 ACGTGACGAGATCTTCCCGTCGCATACCTGGAGAGGTGTAGTCGCGCCAGGTCC 7523
OY 832 TTGCTACACGACGACGACAGTGTGCCCGCAGCGCGCCACCCCGTTATTTGATTTGATC 891
|||
Db 7522 TGGCCGTAGGCGCCACCGCTTGGCGCGCAGTAGCATACGTTGCCATGTACTTGGCC 7463
OY 892 ATGCAGCGT 901
|||
Db 7462 TTGGGTGTGT 7453

Search completed: June 12, 2003, 08:19:34
Job time : 111 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: June 12, 2003, 08:18:44 ; Search time 357 Seconds

(Without alignments)
8345.641 Million cell updates/sec

Title: US-09-777-566a-1

Perfect score: 1323
Sequence: 1 atgaagcgatcttaacc.....atcacatcacatcactaa 1323

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1322	99.9	1323	22	AAC88885
2	1322	99.9	1323	22	Escherichia coli B
3	1322	99.9	1323	24	AAD36473
4	1287.8	97.3	1901	24	AAD25460
5	1287	97.3	1299	21	ABR12514
6	1282.2	96.9	3470	21	AAC68298
7	1282.2	96.9	4060	21	AAC68296
8	1282.2	96.9	6116	21	AAC68297
9	1282.2	96.9	6708	21	AAC68295

10	1282.2	96.9	17732	21	AAC68300	Lam2/APPA plasmid
11	1282.2	96.9	20623	21	AAC68294	Lam2/APPA plasmid
12	1270.2	96.0	5421	21	AAC68299	SV40/APPA plasmid
13	1263.8	95.5	1489	21	AA28216	E. coli acidic phosph
14	1263	95.5	1486	22	AAD06831	E. coli acid phosph
15	1251.8	94.6	1486	22	AAD06832	E. coli acid phosph
16	1210	91.5	1272	20	AA26540	Genomic sequence #
17	820.4	62.0	846	22	AA51940	DNA encoding novel
18	689.4	52.1	1737	23	AA589885	DNA encoding novel
19	482.2	36.4	519	22	AA51139	CDNA encoding novel
20	399.4	30.2	4050	23	AA564562	DNA encoding novel
21	390	29.5	1264	23	AA588443	DNA encoding novel
22	390	29.5	1264	23	AA594316	DNA encoding novel
23	199	15.0	356	23	AA589883	DNA encoding novel
24	157	11.9	5976	23	AA588450	DNA encoding novel
25	56.8	4.3	11710	24	AAH3718	E. coli genome sec
26	39.8	3.0	2868	22	AAD05673	Mouse zmsel cDNA.
27	36.2	2.7	427	21	AA567371	Pinus radiata gluc
28	36	2.7	450	23	ABV05446	Human prostate exp
29	35.4	2.7	344	23	ABV04938	Human prostate exp
30	35.4	2.7	349	23	ABV14107	Human prostate exp
31	35.4	2.7	391	23	ABV35209	Human prostate exp
32	35.4	2.7	391	23	ABV44049	Human prostate exp
33	35.4	2.7	405	23	ABV14615	Human prostate exp
34	35.4	2.7	437	23	ABV35693	Human prostate exp
35	35.4	2.7	2781	23	ABK42477	Genomic sequence #
36	34.4	2.6	800	22	AAH04926	Human cDNA clone (
37	34.4	2.6	3076	22	AAD05671	Human zmsel cDNA.
38	34.4	2.6	3087	22	AA193860	Human stomach canc
39	34.4	2.6	3087	22	AAH17541	Human cDNA sequenc
40	34.4	2.6	4403765	22	AA199683	Mycobacterium tube
41	34.4	2.6	4411529	22	AA199682	Mouse amphiregulin
42	34	2.6	998	19	AAV59100	Human zmsel cDNA,
43	34	2.6	1068	22	AAD05672	Drosophila melanog
44	34	2.6	1613	23	ABL05642	Human prostate exp
45	33.8	2.6	1979	23	ABV20205	

ALIGNMENTS

RESULT 1	
AAC88885	
ID	AAC88885 standard; DNA; 1323 BP.
AC	AAC88885;
XX	
DT	07-MAR-2001 (first entry)
XX	
DE	Escherichia coli B phytase enzyme nucleotide sequence.
KW	Escherichia coli B; phytase enzyme; anabolic; phytate digestion;
KW	nutrition; ds.
OS	Escherichia coli.
XX	
PN	WO200071728-A1.
XX	
PD	30-NOV-2000.
PF	25-MAY-2000; 2000WO-US14846.
XX	
PR	25-MAY-1999; 99US-0318528.
XX	
PA	(DIVE-) DIVERSA CORP.
XX	
PI	Short JM, Kretz KA;
XX	
DR	WPI: 2001-112081/12.
XX	
DR	P-PSDB; AAB37892.
PT	Improving the nutritional value of phytate-containing foodstuffs, using
PT	phytase enzymes which catalyze the liberation of inorganic phosphate

from the phytates -
 Claim 2: Fig 1: 147pp; English.

The present sequence encodes a phytase enzyme from *Escherichia coli* B. The enzyme catalyses the liberation of inorganic phosphate from the phytate in phytate-containing foodstuffs and can thus be used to improve the nutritional value of phytate rich ingredients.

Sequence 1323 BP: 323 A; 353 C; 357 G; 289 T; 1 other:

Query Match 99.9%; Score 1322; DB 22; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCAATCTGA
DB 1 ATGAAGGAGATCTTAATCCATTTTATCTTCTGATTCGTTAACCCCGCAATCTGA
OY 61 TTGCTCAGAGTGAAGCTGGAAGTGTGATTTGTCAGTGCATGCT 120
DB 61 TTGCTCAGAGTGAAGCTGGAAGTGTGATTTGTCAGTGCATGCT 120
OY 121 GTGCGTGTCCCAACCAAGCCGCACTGATGACAGATGTACCCCGAGAGCATGGCA
DB 121 GTGCGTGTCCCAACCAAGCCGCACTGATGACAGATGTACCCCGAGAGCATGGCA
OY 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCCGCGGGTGTGACTAATCCCTATCTC
DB 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCCGCGGGTGTGACTAATCCCTATCTC
OY 241 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGCTGCCG
DB 241 GGACATTACCAACGCGACGCTGTGTAGCCGAGATTTGCTGGCAAAAAGGCTGCCG
OY 301 CAGTGTGCAAGTGTGCGATTTATGCTGATGCGAGCGGATACCGGTAACAGCGGA
DB 301 CAGTGTGCAAGTGTGCGATTTATGCTGATGCGAGCGGATACCGGTAACAGCGGA
OY 361 GCGTTCGCGCGGGCTGGCACTGCTGCAATTAACCGTATACCGAGCAGATAGC
DB 361 GCGTTCGCGCGGGCTGGCACTGCTGCAATTAACCGTATACCGAGCAGATAGC
OY 421 TCCAGTCCCGATCCCTTATTTATCTCTAAAACTGGCGTTGGCAACTGTATACCG
DB 421 TCCAGTCCCGATCCCTTATTTATCTCTAAAACTGGCGTTGGCAACTGTATACCG
OY 481 AACGTGACTGACGGATCTCAGCAGGAGGAGGATTAATGCTGACTTACCGGGCAT
DB 481 AACGTGACTGACGGATCTCAGCAGGAGGAGGATTAATGCTGACTTACCGGGCAT
OY 541 CGGCAAAAGCGGTTTGGCAACTGGAACGGGTGCTTAAATTTTCCGAATCAAACTTGTG
DB 541 CGGCAAAAGCGGTTTGGCAACTGGAACGGGTGCTTAAATTTTCCGAATCAAACTTGTG
OY 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGAGCATTAACCATGCACTC
DB 601 CTTAAACGTGAGAAACAGAGCAAGAGCTGTTCAATTAACGAGCATTAACCATGCACTC
OY 661 AAGGTGAGCGCGGACATGTCTCATTAACCGGTGGGTAAAGCTGCAATCATGCTGACG
DB 661 AAGGTGAGCGCGGACATGTCTCATTAACCGGTGGGTAAAGCTGCAATCATGCTGACG
OY 721 GAGATATTTCTCTGCAACAGCAAGAGGAATGCCGAGCGGGGTGGGAGAGATCAC
DB 721 GAGATATTTCTCTGCAACAGCAAGAGGAATGCCGAGCGGGGTGGGAGAGATCAC
OY 781 GATTACACACAGTGAACACCTGCTAAGTTGATTAACGCGCAATTTATTTGCTACAA
DB 781 GATTACACACAGTGAACACCTGCTAAGTTGATTAACGCGCAATTTATTTGCTACAA
OY 841 CGCAGCGCAGAGTGGCGCGGACCGCGCCGCTTATGATTTGATCATGTCACG
DB 841 CGCAGCGCAGAGTGGCGCGGACCGCGCCGCTTATGATTTGATCATGTCACG

```

```

DB 841 CGCAGCGCAGAGTGGCGCGGACCGCGCCGCTTATGATTTGATCATGTCACG 900
OY 901 TTGAGCGCCCATCCACCCGAAAAACAGGCGTATGGTGTGATTAATCCACTTCAGTAC
DB 901 TTGAGCGCCCATCCACCCGAAAAACAGGCGTATGGTGTGATTAATCCACTTCAGTAC
OY 961 TTTATTTGCGGACAGATTAATCTGCAAAATCTCGGCGGCGCACTGGAGCTCAACTG
DB 961 TTTATTTGCGGACAGATTAATCTGCAAAATCTCGGCGGCGCACTGGAGCTCAACTG
OY 1021 ACGCTTCCCGGTCCAGCCGATTAACACCGCCGAGGTGTGAGACTGTTTAACGCTG
DB 1021 ACGCTTCCCGGTCCAGCCGATTAACACCGCCGAGGTGTGAGACTGTTTAACGCTG
OY 1081 ACGCTTCCCGGTCCAGCCGATTAACACCGCCGAGGTGTGAGACTGTTTAACGCTG
DB 1081 ACGCTTCCCGGTCCAGCCGATTAACACCGCCGAGGTGTGAGACTGTTTAACGCTG
OY 1141 CAGATGCGGTATTAACACCGCGCTGTCAATTAATACCGCGCCGAGAGGTGAATGAC
DB 1141 CAGATGCGGTATTAACACCGCGCTGTCAATTAATACCGCGCCGAGAGGTGAATGAC
OY 1201 CTGGCAGAGTGTGAAGAGCGAATGCGAGGCGATGTGCTGTGGCAGGTTTACGAA
DB 1201 CTGGCAGAGTGTGAAGAGCGAATGCGAGGCGATGTGCTGTGGCAGGTTTACGAA
OY 1261 ATCGTAATGAAGCAGCAGATACCGGCGTGCAGTGTGAGATTCATCACCATCCATC
DB 1261 ATCGTAATGAAGCAGCAGATACCGGCGTGCAGTGTGAGATTCATCACCATCCATC
OY 1321 TAA 1323
DB 1321 TAA 1323

RESULT 2
AAD36473
ID AAD36473 standard; DNA; 1323 BP.
XX
XX AAD36473;
AC 21-AUG-2002 (first entry)
XX
XX Escherichia coli phytase DNA.
DE Dietary aid: biocompatible composition; therapeutic; digestive tract;
XX foodstuff; digestion; phytase; enzyme; gene; ds.
XX
XX Escherichia coli.
OS
FH Key Location/Qualifiers
FT 1..1323
FT CDS /tag="a
FT product="E. coli phytase protein"
XX
XX WO200189317-A2.
XX
XX 29-NOV-2001.
XX
XX 15-MAY-2001; 2001WO-US15764.
XX
XX 25-MAY-2000; 2000US-0580937.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz KA, O'Donoghue E.
XX
XX WPI: 2002-164149/21.
XX
XX P-PSDB: AAE22836.
XX
XX New dietary aids comprising sustained release biocompatible
XX compositions, comprise agent that assists in digestion, useful for
XX delivering enzymes, therapeutics, medicine or agents to an organism
XX

```


XX	Key	Location/Qualifiers
FH	CDS	188..1486
FT		/tag= a
ET		/product= "E. coli appa phytase protein"
FN		
PM	WO200190333-A2.	
PD		
PP	29-NOV-2001.	
PR	24-MAY-2001; 2001WO-US17118.	
PS	25-MAY-2000; 2000US-0580515.	
PX	(DIVE-) DIVERSA CORP.	
RA		
RI	Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;	
RJ	WPI: 2002-083108/11.	
DR	P-PsDB: MAE15807.	
XX		
PT	New bacterial phytase for e.g. improving the nutritional value of	
PF	phytate-containing foodstuffs and subsequently improving the growth	
PE	performance of an organism that consumes it, or in treating animal	
PP	digestive systems -	
PS	Claim 42; Fig 7; 170pp; English.	
XX		
CC	The patent discloses recombinant bacterial phytase from Escherichia coli	
CC	K12 appa phytase. The enzyme has phytase activity and improved thermal	
CC	tolerance when compared with wild-type phytase. It has improved protease	
CC	stability at low pH. The recombinant phytase is useful for improving the	
CC	nutritional value of phytate-containing foodstuffs and subsequently	
CC	improving the growth performance of an organism that consumes it, in	
CC	treating animal digestive systems, in feed treatment processes and for	
CC	in vitro purposes related to research, discovery and development. They	
CC	are also used for generating recombinant digestive system life forms,	
CC	for producing or manufacturing alcoholic and non-alcoholic drinks based	
CC	on the use of moulds, grains and/or plants, in biopulping and bio-	
CC	bleaching where a reduction in the use of environmentally harmful	
CC	chemicals that are traditionally used in the pulp and paper industry	
CC	is desired and in the reduction or possible elimination of the need	
CC	for mineral supplements, enzymes or therapeutic drugs for animals	
CC	from the daily feed thus increasing the amount calories and nutrients	
CC	present in the feed. The present sequence is a DNA encoding E. coli	
CC	appa phytase wild type protein.	
XX		
SQ	Sequence 1901 BP; 474 A; 499 C; 499 G; 428 T; 1 other:	
	Query Match 97.3%; Score 1287.8; DB 24; Length 1901;	
	Best Local Similarity 99.5%; Pred. No. 0;	
	Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
OY	1 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCGTACGTTAACCCGGCAATCTGA 60	
DB	188 ATGAAAGCGATCTTAATCCATTTTTATCTCTTCGTACGTTAACCCGGCAATCTGA 247	
OY	61 TTCCGCTCAGATGAGCCGCGAGACTGAAAGCTGGAAGTGCGGATTTGTAAGTCGTATGCT 120	
DB	248 TTCCGCTCAGATGAGCCGCGAGACTGAAAGTGCGGATTTGTAAGTCGTATGCTATGCT 307	
OY	121 GTGGCGTGTCCAACCAAGGCCACGCACTGTAGTAGCGAGATGTCACCACCGACGATGGGCA 180	
DB	308 GTGGCGTGTCCAACCAAGGCCACGCACTGTAGTAGCGAGATGTCACCACCGACGATGGGCA 367	
OY	181 ACCTGGCCGGTAAAACTGGGTTGGCTGCACACCGCGNGGTGGTGAAGCTAATCGCCCTATCTC 240	
DB	368 ACCTGGCCGGTAAAACTGGGTTGGCTGCACACCGCGNGGTGGTGAAGCTAATCGCCCTATCTC 427	
OY	241 GGACATTACCAAACGCGCACGCTGTGTTAGCCGACGAGATTGCTGGCGAAAAGGCGTCCCG 300	
DB	428 GGACATTACCAAACGCGCACGCTGTGTTAGCCGACGAGATTGCTGGCGAAAAGGCGTCCCG 487	

OY	301	CAGCTGGTCAAGGTCGGGATTAATTTGATGTGACAGGAGCGTACCCGTAATAAAGCGGAA	360
OY	361	GCTTTCGGCGCGGGGGTGGACCTGACATGTGCAATTAACCTGATACCTGACAGGAGATACG	420
OY	488	CAGTCTGGTCAAGTCCGGCATTTATGTCTGATGTGACAGGAGCGGTACCCGTAATAAAGCGGAA	547
OY	361	GCTTTCGGCGCGGGGGTGGACCTGACATGTGCAATTAACCTGATACCTGACAGGAGATACG	420
OY	481	AACGTGACTGACCGGGATCTCTAGACAGGGCAGAGAGGTCAATTGCTGACTTACCGGGCAT	540
OY	608	TCCAGTCCCGGATCCGTTATTAATCTCTAAAACTGGCGTTTGCCAACTGATTAACGCG	667
OY	541	CGGCAAAACGGGGTTCGGCAACTGGAACGGGGTGCCTTAATTTCCGCAATCAAACTTGTC	600
OY	728	CGGCAAAACGGGGTTCGGCAACTGGAACGGGGTGCCTTAATTTCCGCAATCAAACTTGTC	787
OY	601	CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACAGGATTAACATCGGAATC	660
OY	788	CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACAGGATTAACATCGGAATC	847
OY	661	AAGGTGAGCGCCGACATGTCTCATTTAACCGCGTGGGTAACCTTCGCATCAATGCTAGC	720
OY	848	AAGGTGAGCGCCGACATGTCTCATTTAACCGCGTGGGTAACCTTCGCATCAATGCTAGC	907
OY	721	GAGATATTTCCTCGCAACAGCAACAGGAGATGCGGAGCGGGGGTGGGAAGGATACCC	780
OY	908	GAGATATTTCCTCGCAACAGCAACAGGAGATGCGGAGCGGGGGTGGGAAGGATACCC	967
OY	781	GATTACACACCAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA	840
OY	968	GATTACACACCAGTGGAAACACCTTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA	1027
OY	841	CGCACGCCAGAGGTTGGCCGACGCGCGCCACCCCGTATTTGATTTGATATATGCGACGC	900
OY	1028	CGCACGCCAGAGGTTGGCCGACGCGCGCCACCCCGTATTTGATTTGATATATGCGACGC	1087
OY	901	TTGACAGCCCCATCCACGCAAAACAGGCGATGATGATGACATTAACCACTTCAGTACTG	960
OY	1088	TTGACAGCCCCATCCACGCAAAACAGGCGATGATGATGACATTAACCACTTCAGTACTG	1147
OY	961	TTTATTGCGGACAGATACTAATCTGGCAATCTCGCGCGGACCTGAGACTCAACTGG	1020
OY	1148	TTTATTGCGGACAGATACTAATCTGGCAATCTCGCGCGGACCTGAGACTCAACTGG	1207
OY	1021	ACGCTTCCCGGTCAACCGCGGATTAACACGCGCCAGAGTGGTAATCTGTGTTGAACGCTGG	1080
OY	1208	ACGCTTCCCGGTCAACCGCGGATTAACACGCGCCAGAGTGGTAATCTGTGTTGAACGCTGG	1267
OY	1081	CGTGGGCTTAACGATTAACAGCCAGTGGATTAAGTTTCGCGGTCCGATCCAGACTTTACAG	1140
OY	1268	CGTGGGCTTAACGATTAACAGCCAGTGGATTAAGTTTCGCGGTCCGATCCAGACTTTACAG	1327
OY	1141	CAGATGGCTGATTAACAGCCGCTGTCTCAATTAATTAACGCGCCCGGAGAGGTGAACCTGACC	1200
OY	1328	CAGATGGCTGATTAACAGCCGCTGTCTCAATTAATTAACGCGCCCGGAGAGGTGAACCTGACC	1387
OY	1201	CTGGCAGAGATGTAAGAGCAAAATGCGCAGGCGATGTGTCTGTTGGCAGGTTTACGCCAA	1260
OY	1388	CTGGCAGAGATGTAAGAGCAAAATGCGCAGGCGATGTGTCTGTTGGCAGGTTTACGCCAA	1447
OY	1261	ATCGTGAATGAAGACGATACCGCGCGGCTGAGTGGAGAT	1300
OY	1448	ATCGTGAATGAAGACGATACCGCGCGGCTGAGTGGAGAT	1487

```

AC  ABR12514;
XX
DT  05-JUN-2002 (first entry)
DE  DNA encoding phytase associated protein.
XX
KM  Phytase; gene: ds.
XX
OS  unidentified.
XX
FH  Key
FT  CDS
FT  1..1299
FT  /lag= a
FT  /product= "Phytase associated protein"
XX
PN  KR9086028-A.
XX
PD  15-DEC-1999.
XX
PF  25-MAY-1998; 98KR-0018810.
XX
PR  25-MAY-1998; 98KR-0018810.
XX
PA  (MOOI-) MOOJIN CO LTD.
XX
PI  Bae HD, Forceburgh CW, Goloben S, Cheng KJ;
XX
DR  WPI: 2000-645078/62.
XX
XX  P-PSDB: AAU77775.
XX
PT  Novel phytase gene, recombinant phytase and usage thereof -
XX
PS  Claim 1; Fig 2; 10pp; Korean.
XX
CC  The invention relates to a novel phytase gene, a recombinant
CC  phytase gene and their uses. This sequence encodes a phytase
CC  associated protein, described in the invention.
XX
SQ  Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 other;

Query Match      97.3%; Score 1287; DB 21; Length 1299;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1290; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY  1 ATGAAAGCGATCTTAATCCCATTTTATCTTCTGATTCGTTAACCCGCAATCTGCA 60
DB  1 ATGAAAGCGATCTTAATCCCATTTTATCTTCTGATTCGTTAACCCGCAATCTGCA 60
QY  61 TTGCGTCAAGATGAGCGGAGCTGAGCTGGAAGTGTGATTTGTCAGTCTCATGCT 120
DB  61 TTGCGTCAAGATGAGCGGAGCTGAGCTGGAAGTGTGATTTGTCAGTCTCATGCT 120
QY  121 GTGCGTCTCAACCAAGGCGACGCACTGATGAGATGTCACCCAGACGATGGCCA 180
DB  121 GTGCGTCTCAACCAAGGCGACGCACTGATGAGATGTCACCCAGACGATGGCCA 180
QY  181 ACCTGGCGGTAAACCTGGTTGGCTGACACCGCGNGGGTGGTGAATTCGCTATCTC 240
DB  181 ACCTGGCGGTAAACCTGGTTGGCTGACACCGCGNGGGTGGTGAATTCGCTATCTC 240
QY  241 GGACATTACCAACGCCAGCGTCTGTGAGCCGACGAGATTCGCTAAACAGGGGAA 300
DB  241 GGACATTACCAACGCCAGCGTCTGTGAGCCGACGAGATTCGCTAAACAGGGGAA 300
QY  301 CAGTCTGTGAGTGGGATTTATGTCGATGTGACAGAGGTACCCGTAACAGGGGAA 360
DB  301 CAGTCTGTGAGTGGGATTTATGTCGATGTGACAGAGGTACCCGTAACAGGGGAA 360
QY  361 GCCTTCGCGCGGGGCTGGCACTGACTGTGCATTAACCGATACCCAGGCGAGATACG 420
DB  361 GCCTTCGCGCGGGGCTGGCACTGACTGTGCATTAACCGATACCCAGGCGAGATACG 420
QY  421 TCCAGTCCCGATCCGTTATTTATCTCTAATAAACTGGCGTTTGCACAATGGATACGCG 480

```

```

DB  421 TCCAGTCCCGATCCGTTATTTAACTCTTAATAAACTGGCGTTTGCACAATGGATACGCG 480
QY  481 AAGTGAAGCGCGGATCTCAGCAGGCGACGAGGAGTCAATTTGCTTTACGGGCAT 540
DB  481 AAGTGAAGCGCGGATCTCAGCAGGCGACGAGGAGTCAATTTGCTTTACGGGCAT 540
QY  541 CGGCAAAACGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
DB  541 CGGCAAAACGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
QY  601 CTTAAGGTGGAAGAAACGAGAGAAAGCTTTTCAATTAACGAGGATTAACATCGGAATC 660
DB  601 CTTAAGGTGGAAGAAACGAGAGAAAGCTTTTCAATTAACGAGGATTAACATCGGAATC 660
QY  661 AAGTGAAGCGCGGATCTCAGCAGGCGACGAGGAGTCAATTTGCTTTACGGGCAT 720
DB  661 AAGTGAAGCGCGGATCTCAGCAGGCGACGAGGAGTCAATTTGCTTTACGGGCAT 720
QY  721 GAGATATTTCTCTGCAACCAAGCAGAGGAATGCCGAGCGGGGTGGGAAGATGACC 780
DB  721 GAGATATTTCTCTGCAACCAAGCAGAGGAATGCCGAGCGGGGTGGGAAGATGACC 780
QY  781 GATTCAACACAGTGGAAACACCTTGTCTAAGTTTGGATAACGGGCAATTTATTGCTCAAA 840
DB  781 GATTCAACACAGTGGAAACACCTTGTCTAAGTTTGGATAACGGGCAATTTATTGCTCAAA 840
QY  841 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCGTTATGATGATGATGATGATGATGATG 900
DB  841 CGCAGCGCAGAGGTTGCCCGCAGCGCGCCACCCGTTATGATGATGATGATGATGATGATG 900
QY  901 TTGAGCGCCCATTCACCGCAAAACAGCGGTATGATGATGATGATGATGATGATGATGATG 960
DB  901 TTGAGCGCCCATTCACCGCAAAACAGCGGTATGATGATGATGATGATGATGATGATGATG 960
QY  961 TTTATTCGCGGACACGATCTAATCTGCAAAATCTGCGGCGGACGCTGAGACCTCACTGG 1020
DB  961 TTTATTCGCGGACACGATCTAATCTGCAAAATCTGCGGCGGACGCTGAGACCTCACTGG 1020
QY  1021 ACCTTCGCGGTCAGCGGATTAACAGCGCGCGGATGATGATGATGATGATGATGATGATGATG 1080
DB  1021 ACCTTCGCGGTCAGCGGATTAACAGCGCGCGGATGATGATGATGATGATGATGATGATGATG 1080
QY  1081 CGTGGGCTAAGCGATTAACAGCGCACTGATGATGATGATGATGATGATGATGATGATGATG 1140
DB  1081 CGTGGGCTAAGCGATTAACAGCGCACTGATGATGATGATGATGATGATGATGATGATGATG 1140
QY  1141 CAGATGCGGTAAACAGCGCGCTGCTAATTAATACGCGCGCGGAGAGGTAAACTGACC 1200
DB  1141 CAGATGCGGTAAACAGCGCGCTGCTAATTAATACGCGCGCGGAGAGGTAAACTGACC 1200
QY  1201 CTGGCAGATGTGAAGAGCGGAATGCGCAGGCGATGTTGCTGGCAGGTTTACGCAA 1260
DB  1201 CTGGCAGATGTGAAGAGCGGAATGCGCAGGCGATGTTGCTGGCAGGTTTACGCAA 1260
QY  1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296
DB  1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTG 1296

RESULT 6
AAC68298
ID  AAC68298 standard; DNA; 3470 BP.
XX  AAC68298;
XX  AC
XX  20-FEB-2001 (first entry)
DT  20-FEB-2001 (first entry)
XX  R15/APPA plasmid coding sequence.
DE  Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KM  environmental pollution; pig; ds.
XX

```

OS Ratus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX PN MO200064247-A1.
 XX PD 02-NOV-2000.
 XX PF 20-APR-2000; 2000MO-CA00430.
 XX PR 23-APR-1999; 99US-0130508.
 XX PA (UYGU-) UNIV GUELPH.
 PI Forsberg CW, Golovan S, Phillips JP;
 XX WPI; 2000-687245/67.
 DR P-PDB; AAB36261.
 XX PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein.
 XX PS Claim 14; Fig 21; 152pp; English.
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APFA phytase coding sequence.
 SO Sequence 3470 BP; 1065 A; 721 C; 735 G; 949 T; 0 other;

Query Match 96.98; Score 1282.2; DB 21; Length 3470;
 Best Local Similarity 99.38; Pred. No. 0;
 Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

QY 1. ATGAAGGATCTTATCCATTTTATCTCTGATTCCTTAACCCCGCAATCGCA 60
 DB 1811 ATGAAGGATCTTATCCATTTTATCTCTGATTCCTTAACCCCGCAATCGCA 1870
 QY 61 TTGCGTCAGAGTGGAGCGGAGTGAAGTGAAGTGGTGAATTCAGTCTCATGT 120
 DB 1871 TTGCGTCAGAGTGGAGCGGAGTGAAGTGAAGTGGTGAATTCAGTCTCATGT 1930
 QY 121 GTGCGTGTCCAAACCAAGGCCACGACTGATGACAGATGTCAACCCAGAGCATGGCCA 180
 DB 1931 GTGCGTGTCCAAACCAAGGCCACGACTGATGACAGATGTCAACCCAGAGCATGGCCA 1990
 QY 181 ACCTGGCCGGTAAACTGGGTGTGACACCGCGGAGTGTGAGTGAATTCGCTATCTC 240
 DB 1991 ACCTGGCCGGTAAACTGGGTGTGACACCGCGGAGTGTGAGTGAATTCGCTATCTC 2050
 QY 241 GGACATTACCAAGCCAGCGTCTGTAGCCGAGATGGTGGCGAAGGAGGCTGCCG 300
 DB 2051 GGACATTACCAAGCCAGCGTCTGTAGCCGAGATGGTGGCGAAGGAGGCTGCCG 2110
 QY 301 CAGTGTGTGAGTGTGATTTATGCTGATGTGACAGAGATGTCAACCCGTAACAGAGCAA 360
 DB 2111 CAGTGTGTGAGTGTGATTTATGCTGATGTGACAGAGATGTCAACCCGTAACAGAGCAA 2170
 QY 361 GCGTTGGCGCGCGGGTGGACACTGTGCAATAACGATACATCCAGGAGATACG 420
 DB 2171 GCGTTGGCGCGCGGGTGGACACTGTGCAATAACGATACATCCAGGAGATACG 2230
 QY 421 TCCAGTCCCGATCCGTTATTTATCTCTTAAATAAGGCGTTGGCAACGATAACGCG 480
 DB 2231 TCCAGTCCCGATCCGTTATTTATCTCTTAAATAAGGCGTTGGCAACGATAACGCG 2290
 QY 481 AACGTGACTGACGAGTCTCAGACAGGAGGATGCAATTCGACTTACCGGAGCAT 540
 DB 2291 AACGTGACTGACGAGTCTCAGACAGGAGGATGCAATTCGACTTACCGGAGCAT 2350

QY 541 CGGCAAAAGGCGGTTTGGCGAAGTGAACGGGTGCTTAATTTCCGAATCAACTGTGC 600
 DB 2351 CGGCAAAAGGCGGTTTGGCGAAGTGAACGGGTGCTTAATTTCCGAATCAACTGTGC 2410
 QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGAGCATTTACATCGAATC 660
 DB 2411 CTTAAACGTGAGAAACAGACGAAAGCTGTTCATTAAACGAGCATTTACATCGAATC 2470
 QY 661 AAGGTGAGCGCGACAAATGTCTATTAAACCGGTGCGGTAAAGCTCGATCATGTGACG 720
 DB 2471 AAGGTGAGCGCGACAAATGTCTATTAAACCGGTGCGGTAAAGCTCGATCATGTGACG 2530
 QY 721 GAGATATTTCTCTGCAACAGACAGGGAATGCGGAGCGGGGGGGAAGATATCAC 780
 DB 2531 GAGATATTTCTCTGCAACAGACAGGGAATGCGGAGCGGGGGGGAAGATATCAC 2590
 QY 781 GATTACACACAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 840
 DB 2591 GATTACACACAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 2650
 QY 841 CGCAGCGCAGAGGTTGCCCGCAGCGCGGCCACCCGTTATTTGATTCATGACGCG 900
 DB 2651 CGCAGCGCAGAGGTTGCCCGCAGCGCGGCCACCCGTTATTTGATTCATGACAGCGC 2710
 QY 901 TTGACGCGCCATCCACCGCAAAACAGGCGGTATGTTGATTAACCCACTTACATAC 960
 DB 2711 TTGACGCGCCATCCACCGCAAAACAGGCGGTATGTTGATTAACCCACTTACATAC 2770
 QY 961 TTTATTTGCCGACACGATTAATCTGCAAAATCTGCGCGGCGCAGTGAAGTCAACTGG 1020
 DB 2771 TTTATTTGCCGACACGATTAATCTGCAAAATCTGCGCGGCGCAGTGAAGTCAACTGG 2830
 QY 1021 ACCTTCCGCGGTAGCCGATTAACAGCGCCGCAAGTGTGATGTTGTTGAACGCTGG 1080
 DB 2831 ACCTTCCGCGGTAGCCGATTAACAGCGCCGCAAGTGTGATGTTGTTGAACGCTGG 2890
 QY 1081 CGTGGCGTAAGCATTAACAGCGAGTGTGATGTTGCTGCTTCCAGACTTTTACAG 1140
 DB 2891 CGTGGCGTAAGCATTAACAGCGAGTGTGATGTTGCTGCTTCCAGACTTTTACAG 2950
 QY 1141 CAGATGCGTATTAACCGCGCTGTCAATTAATACCGCGCGGAGAGTGAACATGAC 1200
 DB 2951 CAGATGCGTATTAACCGCGCTGTCAATTAATACCGCGCGGAGAGTGAACATGAC 3010
 QY 1201 CTGGCAGAGTGTGAAGCGCAATTCGCGAGGCGATGTCTTGTGGAGTTTACGCAA 1260
 DB 3011 CTGGCAGAGTGTGAAGCGCAATTCGCGAGGCGATGTCTTGTGGAGTTTACGCAA 3070
 QY 1261 ATGCTGAATGAACGACGATACCGGCGTGTGAGTTTG 1296
 DB 3071 ATGCTGAATGAACGACGATACCGGCGTGTGAGTTTG 3106

RESULT 7
 AAC68296
 ID AAC68296 standard; DNA; 4060 BP.
 XX AAC68296;
 AC 20-FEB-2001 (first entry)
 DE R15/APFA plasmid coding sequence.
 XX Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KW environmental pollution; pig; ds.
 OS Ratus sp - chimeric.
 OS Escherichia coli - chimeric.
 XX PN MO200064247-A1.
 XX PD 02-NOV-2000.

PF	20-APR-2000; 2000OWO-CA00430.
XX	23-APR-1999; 990S-0130508.
PR	(UYGU-) UNIV GUELPH.
PA	Forsberg CM, Golovan S, Phillips JP;
XX	WPI: 2000-687245/67.
DR	P-PDB: AAB36239.
XX	
PT	Transgenic non-human animal for gastrointestinal tract specific
CC	expression of a protein, preferably phytase, comprises a nucleic acid
CC	sequence including a heterologous transgene construct encoding the
CC	protein -
PS	Claim 14; Fig 19; 152pp; English.
XX	
CC	The present invention provides transgenic animals which produce desired
CC	proteins, in this case pigs which expresses phytase in the salivary
CC	gland. Low phytase production levels result in phytate in the diet being
CC	excreted and causing phosphorus contamination in water, as well as
CC	reducing the growth of animals. The invention provides a number of
CC	transgenes containing the E. coli APPA phytase coding sequence.
XX	
SQ	Sequence 4060 BP; 1257 A; 814 C; 843 G; 1146 T; 0 other;
Query Match	96.9%; Score 1282.2; DB 21; Length 4060;
Best Local Similarity	99.3%; Pred. No. 0;
Matches 1287; Conservative.	0; Mismatches 9; Indels 0; Gaps 0;
OY	1 ATGAAGCGATCTTAATCCATTTCCTTCTGATTCGGTTAACCCGCACATGTGCA 60
Db	1811 ATGAAMCCCATCTTAATCCCATTTTATCTCTGATTCGGTTAACCCGCACATGTGCA 1870
OY	61 TTGCTCAGATGAGGCCGCGAGCTAAGCTGSAAGTGCGGATTCGACGTCATGCT 120
Db	1871 TTGCTCAGATGAGGCCGCGAGCTAAGCTGSAAGTGCGGATTCGACGTCATGCT 1930
OY	121 GTGGCTCTCCAAACCAAGGCCACGCAACTGATGAGATGTCACCACAGACATGGCCA 180
Db	1931 GTGGCTCTCCAAACCAAGGCCACGCAACTGATGAGATGTCACCACAGACATGGCCA 1990
OY	181 ACCGAGCCGGTAAACACTGGTGTGCTGACACACCGGNGTGGTGAGCTAATGCGCTATCTC 240
Db	1991 ACCGAGCCGGTAAACACTGGTGTGCTGACACACCGGNGTGGTGAGCTAATGCGCTATCTC 2050
OY	241 GGACATTACCAAACGCCACGCTGTGTAGCCACGACGATTTGGTGGCAAAAAGGCTGCCG 300
Db	2051 GGACATTACCAAACGCCACGCTGTGTAGCCACGACGATTTGGTGGCAAAAAGGCTGCCG 2110
OY	301 CAGTCTGTGTCAAGTTCGGGATTTATGTCGATGTGCAACGAGCGTACCCTGATAACAGCGAA 360
Db	2111 CAGTCTGTGTCAAGTTCGGGATTTATGTCGATGTGCAACGAGCGTACCCTGATAACAGCGAA 2170
OY	361 GCCTTCCCGCGGGGCTGGCACCTGACTGTGCATTAACCGTACAATACAGGACGATACG 420
Db	2171 GCCTTCCCGCGGGGCTGGCACCTGACTGTGCATTAACCGTACAATACAGGACGATACG 2230
OY	421 TTCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTCCCAACTGGATTAACGCG 480
Db	2231 TTCAGTCCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTCCCAACTGGATTAACGCG 2290
OY	481 AACGTGACTGACGGGATCTCAGAGAGGCGAGGAGGTCAAATGTGACTTAACCGGAT 540
Db	2291 AACGTGACTGACGGGATCTCAGAGAGGCGAGGAGGTCAAATGTGACTTAACCGGAT 2350
OY	541 CGGCAAAACGGCGTTCCGCAACTGCAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 600
Db	2351 CGGCAAAACGGCGTTCCGCAACTGCAACGGGTGCTTAATTTCCGCAATCAAACTTGTGC 2410
OY	601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGCAATTCATCGGAATCTC 660

Db	2411	CTTAAAGCTGGAAGAAAGGACGAAAGCGTTTCATTAAAGCAGGACATTACCATCGGAATTC	2470
OY	661	AAGGTGAGCGCCGCAATATGTCATTAAACGGGTGCGGTAAAGCTTCGCATCAATGCTGACG	720
Db	2471	AAGGTGAGCGCCGCAATATGTCATTAAACGGGTGCGGTAAAGCTTCGCATCAATGCTGACG	2530
OY	721	GAGATATTTTCTCCGCAACAAGCAGCAGAGGAATGCGGAGCCGGGGTGGGGAAGATCACC	780
Db	2531	GAGATATTTTCTCCGCAACAAGCAGCAGAGGAATGCGGAGCCGGGGTGGGGAAGATCACC	2590
OY	781	GATTACACACAGTAGTGAACACCTTGTCTAAGTTTGCATAAACGCCAATTTATTTCTACAA	840
Db	2591	GATTACACACAGTAGTGAACACCTTGTCTAAGTTTGCATAAACGCCAATTTATTTCTACAA	2650
OY	841	CGCAGCGCAGAGGTGGCCCGCAGCGCGGCGACCCCGTATTGATGGATTGATGATGGACGG	900
Db	2651	CGCAGCGCAGAGGTGGCCCGCAGCGCGGCGACCCCGTATTGATGGATTGATGATGGACGG	2710
OY	901	TTGACGCCCATCCACCGCAAAAACAGGCGTATGTTGATTAACCATTTCCAGTACGTG	960
Db	2711	TTGACGCCCATCCACCGCAAAAACAGGCGTATGTTGATTAACCATTTCCAGTACGTG	2770
OY	961	TTTATTTGCGGACACGATCTAATCTGGCAAAATCTCGGCGGCGCAGCTGAGACTCAACTGG	1020
Db	2771	TTTATTTGCGGACACGATCTAATCTGGCAAAATCTCGGCGGCGCAGCTGAGACTCAACTGG	2830
OY	1021	ACGCTTCCCGGTACGCCGATTAACAGCCGCCGACAGTGGTGAACAGTGTTTGAACGCTGG	1080
Db	2831	ACGCTTCCCGGTACGCCGATTAACAGCCGCCGACAGTGGTGAACAGTGTTTGAACGCTGG	2890
OY	1081	CGTGGGCTAAGCGATTAACAGCCAGTGTGATTTGCTGCTGCTTCACAGCTTTTACAG	1140
Db	2891	CGTGGGCTAAGCGATTAACAGCCAGTGTGATTTGCTGCTGCTTCACAGCTTTTACAG	2950
OY	1141	CAGATGGGTGTTAAAGCGCGCTCTCAATTAATTAACGCGCCCGGAGAGAGTGAACCTGACC	1200
Db	2951	CAGATGGGTGTTAAAGCGCGCTCTCAATTAATTAACGCGCCCGGAGAGAGTGAACCTGACC	3010
OY	1201	CTGGCAGAGTAGTGAAGAGCGGAATGCGCAGGCGCATGTGTTGTTGCGCAGGTTTACGCA	1260
Db	3011	CTGGCAGAGTAGTGAAGAGCGGAATGCGCAGGCGCATGTGTTGTTGCGCAGGTTTACGCA	3070
OY	1261	ATCGTGAATGAAGCAGCATTAACGGCGGCGGACGTTTG	1296
Db	3071	ATCGTGAATGAAGCAGCATTAACGGCGGCGGACGTTTG	3106
RESULT 8			
AAC68297 standard; DNA; 6116 BP.			
ID	AAC68297		
XX	AAC68297;		
AC	20-FEB-2001 (first entry)		
XX			
DT			
XX			
XX			
DE	RI5/ABPA plasmid coding sequence.		
XX			
KW	Transgenic animal; salivary protein; phytase; phosphorus; animal growth;		
KW	environmental pollution; pig; ds.		
OS	Rattus sp - chimeric.		
OS	Escherichia coli - chimeric.		
XX			
PN	MO200064247-A1.		
XX			
PD	02-NOV-2000.		
XX			
PF	20-APR-2000; 2000WO-CA00430.		
XX			
PR	23-APR-1999; 99US-0130508.		
XX			
PA	(UTG-) UNIV GUELPH.		

XX
PI Forsberg CW, Golovan S, Phillips JP;
XX
DR WPI: 2000-687245/67.
DR P-PSDB: AAB36250.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein
XX
PS Claim 56; Fig 20; 152pp; English.
XX
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
SO Sequence 6116 BP; 1724 A; 1386 C; 1407 G; 1599 T; 0 other;

Query Match 96.9%; Score 1282.2; DB 21; Length 6116;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

1 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCCGCAATCGCA 60
1811 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCCGCAATCGCA 1870

61 TTGCGTCAGAGTACGCGAGCTGAAGCTGGAAGTGTGATTTGCTGATGCTGATGCT 120
1871 TTGCGTCAGAGTACGCGAGCTGAAGCTGGAAGTGTGATTTGCTGATGCTGATGCT 1930

121 GTGCGTGTCTCAACCAAGCCAGCACTGATGATGATGATGATGATGATGATGATGAT 180
1931 GTGCGTGTCTCAACCAAGCCAGCACTGATGATGATGATGATGATGATGATGATGAT 1990

181 ACCTGGCGGTAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 240
1991 ACCTGGCGGTAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2050

241 GGACATTACCAACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 300
2051 GGACATTACCAACGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2110

301 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
2111 CAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2170

361 GCCTTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 420
2171 GCCTTGGCGCGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2230

421 TCCAGTCCGCGATCGCTTATTAATCTCTAAACGCTGCTGCTGCTGCTGCTGCTGCTG 480
2231 TCCAGTCCGCGATCGCTTATTAATCTCTAAACGCTGCTGCTGCTGCTGCTGCTGCTG 2290

481 AACGTGACTGACGCGATCTCTACAGGCGAGAGGCTCAATGCTGACTTAACCGGCGAT 540
2291 AACGTGACTGACGCGATCTCTACAGGCGAGAGGCTCAATGCTGACTTAACCGGCGAT 2350

541 CGGCAAAACGCGCTTGGCGAAGCTGCTTCAATTAACGAGGATTAACGAGGATTAAC 600
2351 CGGCAAAACGCGCTTGGCGAAGCTGCTTCAATTAACGAGGATTAACGAGGATTAAC 2410

601 CTTAAACGCTGAGAAACGAGCAAGCTGCTTCAATTAACGAGGATTAACGAGGATTAAC 660
2411 CTTAAACGCTGAGAAACGAGCAAGCTGCTTCAATTAACGAGGATTAACGAGGATTAAC 2470

661 AAGGTGAGCGCGCAATGCTCTCAATTAACGAGGATTAACGAGGATTAACGAGGATTAAC 720
2471 AAGGTGAGCGCGCAATGCTCTCAATTAACGAGGATTAACGAGGATTAACGAGGATTAAC 2530

QY 721 GAGATATTTCTCTGCAACAGCAGAGGATGCGGAGCGGGGTGGGAGATCAC 780
DB 2531 GAGATATTTCTCTGCAACAGCAGAGGATGCGGAGCGGGGTGGGAGATCAC 2590

QY 781 GATTACACACAGTGAACACCTTGTCTAATGCTTCAATACGCGCAATTTATTTGTACAA 840
DB 2591 GATTACACACAGTGAACACCTTGTCTAATGCTTCAATACGCGCAATTTATTTGTACAA 2650

QY 841 CGCAGCGCAGAGTGGCG 900
DB 2651 CGCAGCGCAGAGTGGCG 2710

QY 901 TTGACGCGCCATCCACCGCAAAACAGCGCGCTATGCTGATTAACCATTCAGTACTG 960
DB 2711 TTGACGCGCCATCCACCGCAAAACAGCGCGCTATGCTGATTAACCATTCAGTACTG 2770

QY 961 TTTATTTGCGCGACAGTACTAATGCTGCAAAATCTGCGCGCGCGCGCGCGCGCGCGCG 1020
DB 2771 TTTATTTGCGCGACAGTACTAATGCTGCAAAATCTGCGCGCGCGCGCGCGCGCGCG 2830

QY 1021 ACCTTCCCGCTCAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
DB 2831 ACCTTCCCGCTCAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2890

QY 1081 CGTCGCGTAAAGCAGTAAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCG 1140
DB 2891 CGTCGCGTAAAGCAGTAAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCG 2950

QY 1141 CAGATGCGTATTAACAGCGCGGCTGCTAATTAATACGCGCGCGCGCGCGCGCGCGCG 1200
DB 2951 CAGATGCGTATTAACAGCGCGGCTGCTAATTAATACGCGCGCGCGCGCGCGCGCGCG 3010

QY 1201 CTGCGAGAGTGTGAAGCGCAAAATGCGCAGGAGTGTGCTGCTGCTGCTGCTGCTG 1260
DB 3011 CTGCGAGAGTGTGAAGCGCAAAATGCGCAGGAGTGTGCTGCTGCTGCTGCTGCTG 3070

QY 1261 ATCGTAATGAGCAGCAGCATACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1296
DB 3071 ATCGTAATGAGCAGCAGCATACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3106

RESULT 9
AAC68295
ID AAC68295 standard; DNA: 6708 BP.
XX
AC AAC68295;
XX
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid coding sequence.
XX
KW Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW environmental pollution; pig; ds.
XX
OS Rattus sp. - chimeric.
XX Escherichia coli - chimeric.
XX
PN W0200064247-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000WO-CA00430.
XX
PR 23-APR-1999; 99US-0130508.
XX
PA (UVG-) UNIV GUELPH.
XX
PI Forsberg CW, Golovan S, Phillips JP;
XX
XX WPI: 2000-687245/67.
XX
XX P-PSDB: AAB36258.

transgenes containing the E. coli APPA phytase coding sequence.

Sequence 20623 BP; 5449 A; 4847 C; 4902 G; 5424 T; 1 other;

Query Match 96.9%; Score 1282.2; DB 21; Length 20623;

Best Local Similarity 99.3%; Pred. No. 0;

Matches 1287; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

```

QY 1 ATGAAGGATCTTAAATCCATTTTATCTCTGATTCGGTTAAACCCGCAATCTCA 60
DB 12653 ATGAAGGATCTTAAATCCATTTTATCTCTGATTCGGTTAAACCCGCAATCTCA 12712
QY 61 TTCGCTCAGAGTACGCGGAGCTGAGTGAAGTGTGTGATTTGTCAGTCTCAGT 120
DB 12713 TTCGCTCAGAGTACGCGGAGCTGAGTGAAGTGTGTGATTTGTCAGTCTCAGT 12772
QY 121 GTGGGTGCTCCACCAAGGCGCAGCACTGATGAGATGTACCCCGAGAGCATGGCCA 180
DB 12773 GTGGGTGCTCCACCAAGGCGCAGCACTGATGAGATGTACCCCGAGAGCATGGCCA 12832
QY 181 ACCTGGCCGCTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATCGCTATCTC 240
DB 12833 ACCTGGCCGCTAAACTGGGTTGGCTGACACCGGNGGTGGTGAAGCTAATCGCTATCTC 12892
QY 241 GGACATTACCAACGCCACGCTCTGTAGCCGACGAGATTGCTGGCGGAAAAAGGCTGCCG 300
DB 12893 GGACATTACCAACGCCACGCTCTGTAGCCGACGAGATTGCTGGCGGAAAAAGGCTGCCG 12952
QY 301 CAGTCTGCTCAGTACGCGGATATATGCTGATGTGACGAGCGGTACCCGTAACAGGCGGA 360
DB 12953 CAGTCTGCTCAGTACGCGGATATATGCTGATGTGACGAGCGGTACCCGTAACAGGCGGA 13012
QY 361 GCGTTCCGCGCGGCGCTGGCACTGACTGTCAATTAACCGTACATACCCAGGACAGATACG 420
DB 13013 GCGTTCCGCGCGGCGCTGGCACTGACTGTCAATTAACCGTACATACCCAGGACAGATACG 13072
QY 421 TCCAGTCCCGGATCCGTTATTTAATCTCTAATAAACTGGCGTTGGCACTGATTAACCGG 480
DB 13073 TCCAGTCCCGGATCCGTTATTTAATCTCTAATAAACTGGCGTTGGCACTGATTAACCGG 13132
QY 481 AACGTGACTGACGAGATCTCTCAGCAGGAGGAGGATCAATTCGATTAACCGGCGCAT 540
DB 13133 AACGTGACTGACGAGATCTCTCAGCAGGAGGAGGATCAATTCGATTAACCGGCGCAT 13192
QY 541 CGGCAAAACGGCGTTTCGGCACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 600
DB 13193 CGGCAAAACGGCGTTTCGGCACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC 13252
QY 601 CTTAAAGCTGAGAAACAGAGAGAAAGCTGTTCATTAACGAGGCAATTACCATCGGAATCTC 660
DB 13253 CTTAAAGCTGAGAAACAGAGAGAAAGCTGTTCATTAACGAGGCAATTACCATCGGAATCTC 13312
QY 661 AAGGTGAGCGCGCAATATGCTCATTAACCGGTGGGTAACTCGCATGATGTCAG 720
DB 13313 AAGGTGAGCGCGCAATATGCTCATTAACCGGTGGGTAACTCGCATGATGTCAG 13372
QY 721 GAGATATTTTCTCTGCAACAAGACAGAGGAATGCCGAGCCGGGTGGGGAAGATCAACC 780
DB 13373 GAGATATTTTCTCTGCAACAAGACAGAGGAATGCCGAGCCGGGTGGGGAAGATCAACC 13432
QY 781 GATTCACACACAGTGGAAACACTTGTCTAATTAACGGCGCAATTTTATTTGCTACAA 840
DB 13433 GATTCACACACAGTGGAAACACTTGTCTAATTAACGGCGCAATTTTATTTGCTACAA 13492
QY 841 CGCAGCGCAGAGGTGGCGCGCAGCGCGCCCGGTATTTGATTTGATATGATGCGACG 900
DB 13493 CGCAGCGCAGAGGTGGCGCGCAGCGCGCCCGGTATTTGATTTGATATGATGCGACG 13552
QY 901 TTGAGCGCCCATTCACGCAAAACAGCGATGATGATGATGATTAACCATCTCACTAGTAC 960
DB 13553 TTGAGCGCCCATTCACGCAAAACAGCGATGATGATGATGATTAACCATCTCACTAGTAC 13612
QY 961 TTATATTCGCGGACAGATTAATCTGCGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 1020

```

```

DB 13613 TTATATTCGCGGACAGATTAATCTGCGCAAAATCTCGCGCGGCACTGAGACTCAACTGG 13672
QY 1021 ACGGTTCCCGGTCAGCGGATTAACAGCGCGGAGGTGTAAGTGTGTTGAACGCTGG 1080
DB 13673 ACGGTTCCCGGTCAGCGGATTAACAGCGCGGAGGTGTAAGTGTGTTGAACGCTGG 13732
QY 1081 CGTCGGCTAACGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCACACTTTACAG 1140
DB 13733 CGTCGGCTAACGATTAACAGCGAGTGAATTCAGTTTCGCTGCTTCACACTTTACAG 13792
QY 1141 CAGATGCTGATTAACCGCGCTGCTATTAATFACGCGCGCGGAGAGGTAAACTGACC 1200
DB 13793 CAGATGCTGATTAACCGCGCTGCTATTAATFACGCGCGCGGAGAGGTAAACTGACC 13852
QY 1201 CTGGCAGATGTAAGAGACGAAATGCGCAGGAGCATGTGCTGTTGGCAGGTTTACGCA 1260
DB 13853 CTGGCAGATGTAAGAGACGAAATGCGCAGGAGCATGTGCTGTTGGCAGGTTTACGCA 13912
QY 1261 ATCGTAATGAAGCAGCATACCGCGCTGACAGTTTG 1296
DB 13913 ATCGTAATGAAGCAGCATACCGCGCTGACAGTTTG 13948

```

RESULT 12

AAC68299 standard; DNA; 5421 BP.

AAC68299;

20-FEB-2001 (first entry)

SV40/APPA plasmid coding sequence.

Transgenic animal; salivary protein; phytase; phosphorus; animal growth; environmental pollution; pig; ds.

Rhesus macaque polyoma virus - chimeric.

Escherichia coli - chimeric.

WO200064247-A1.

02-NOV-2000.

20-APR-2000; 2000WO-CA00430.

23-APR-1999; 99US-0130508.

(UIGU-) UNIV GUELPH.

Forsberg CW, Golovan S, Phillips JP;

WPI: 2000-687245/67.

P-PSDB: AAB36262.

Transgenic non-human animal for gastrointestinal tract specific expression of a protein, preferably phytase, comprises a nucleic acid sequence including a heterologous transgene construct encoding the protein.

Claim 56; Fig 22; 152pp; English.

The present invention provides transgenic animals which produce desired proteins, in this case pigs which expresses phytase in the salivary gland. Low phytase production levels result in phytate in the diet being excreted and causing phosphorus contamination in water, as well as reducing the growth of animals. The invention provides a number of transgenes containing the E. coli APPA phytase coding sequence.

Sequence 5421 BP; 1413 A; 1321 C; 1331 G; 1355 T; 1 other;

Query Match 96.0%; Score 1270.2; DB 21; Length 5421;

Best Local Similarity 99.3%; Pred. No. 0;

	Matches	1287	Conservative	0	Mismatches	8	Indels	1	Gaps	1
QY	1	ATGAAGCGATCTTAATCCCATTTTATCTTTGTGATTCGGTTAACCCGGCAATCTGCA	60							
Db	40	ATGAAGACCACTTAATCCCATTTTATCTTCTGTGATTCGGTTAACCCGGCAATCTGCA	99							
QY	61	TTGCTCAGAGTAGGCGGAGCTCAGCTGGAAAGTGAGGAGATTGTCAGTCGTCATGCT	120							
Db	100	TTGCTCAGAGTAGGCGGAGCTCAGCTGGAAAGTGAGGAGATTGTCAGTCGTCATGCT	159							
QY	121	GTGCGTCCTCCAAACCAAGGCGCAGCAACTGATGCAAGATGTCACCCCAAGACGATGGCCA	180							
Db	160	GTGCGTCCTCCAAACCAAGGCGCAGCAACTGATGCAAGATGTCACCCCAAGACGATGGCCA	219							
QY	181	ACCTGGCCGGTAAACTGEGTTGGCTGACACCGCGNGGTGGTAGCTAATTCCTATCTC	240							
Db	220	ACCTGGCCGGTAAACTGEGTTGGCTGACACCGCGNGGTGGTAGCTAATTCCTATCTC	279							
QY	241	GGACATTAACCAAGCGCCAGCGTGTGGTGGCCAGCGGATTTGTTGGGAAAAAGGGCTGCCG	300							
Db	280	GGACATTAACCAAGCGCCAGCGTGTGGTGGCCAGCGGATTTGTTGGGAAAAAGGGCTGCCG	339							
QY	301	CAGTCGTGTGACGTCGCGGATTTATGTTCGATGTCACGAGCGTACCCTGGTAAACAGGCGAA	360							
Db	340	CAGTCGTGTGACGTCGCGGATTTATGTTCGATGTCACGAGCGTACCCTGGTAAACAGGCGAA	399							
QY	361	GCCTTGCGCGCGGGGTGGCACCTGACTGTGCAATTAACCGTACATACCAGGCAAGATACG	420							
Db	400	GCCTTGCGCGCGGGGTGGCACCTGACTGTGCAATTAACCGTACATACCAGGCAAGATACG	459							
QY	421	TTCCAGTCCCGGATCCGTTATTTAATTCCTTAATAAACTGGCGTTTGGCCAATCGATACGGC	480							
Db	460	TTCCAGTCCCGGATCCGTTATTTAATTCCTTAATAAACTGGCGTTTGGCCAATCGATACGGC	519							
QY	481	AACGTGACTGACGGGATCCTCAGCAGGGCAGAGAGGTCATTTGCTGACTTACCGGGCAT	540							
Db	520	AACGTGACTGACGGGATCCTCAGCAGGGCAGAGAGGTCATTTGCTGACTTACCGGGCAT	579							
QY	541	CGGCAAAAGCGGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC	600							
Db	580	CGGCAAAAGCGGTTTCGCGAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC	639							
QY	601	CTTAAAGCTGAGAAACAGGAAAGCTGTTCATTAAACGCAAGGCAATTACCTCGGAATCTC	660							
Db	640	CTTAAAGCTGAGAAACAGGAAAGCTGTTCATTAAACGCAAGGCAATTACCTCGGAATCTC	699							
QY	661	AAGGTGAGCGCGCAATGTCTCATTTAAACGGGTGCGGTAACCTCGCATCAATCTGACG	720							
Db	700	AAGGTGAGCGCGCAATGTCTCATTTAAACGGGTGCGGTAACCTCGCATCAATCTGACG	759							
QY	721	GAGATATTTTCTCTGCAACAAGCAGGAGATGCGGAGCGGSGTGGGGAAGATCACG	780							
Db	760	GAGATATTTTCTCTGCAACAAGCAGGAGATGCGGAGCGGSGTGGGGAAGATCACG	819							
QY	781	GATTCAACACAGTGGAAACACCTTGCTAAGTTTGCATTAACGCGCAATTTATTTGCTACAA	840							
Db	820	GATTCAACACAGTGGAAACACCTTGCTAAGTTTGCATTAACGCGCAATTTATTTGCTACAA	879							
QY	841	CGCAGCGCAGAGGTGGCCCGCAGCGCGGCAACCCGTTATTTGAGTTGATCATGAGCGCG	900							
Db	880	CGCAGCGCAGAGGTGGCCCGCAGCGCGGCAACCCGTTATTTGAGTTGATCATGAGCGCG	939							
QY	901	TTGACGCGCCATTCACCGCAAAAACAGGCGCTATGCTGACATTAACCACTTCAGTACTG	960							
Db	940	TTGACGCGCCCA - CCACGGCAAAAACAGGCGTATGGTGTGACATTAACCACTTCAGTGTG	998							
QY	961	TTTATTGCGGACACGATTAATCTGGCAAAATCTGGCGGCGCACTGAGACTCAACTGG	1020							
Db	999	TTTATTGCGGACACGATTAATCTGGCAAAATCTGGCGGCGCACTGAGACTCAACTGG	1058							
QY	1021	ACGGTTCCCGGTGACGGGATTAACAGCGCCGAGGTGTGAACGTGGTTTGAACGCTGG	1080							
Db	1059	ACGGTTCCCGGTGACGGGATTAACAGCGCCGAGGTGTGAACGTGGTTTGAACGCTGG	1118							

QY	1081	CGTGGCGTAAAGCCGTAACAGCCAGTGAATTCAGGTTTCGGTGGCTTCAGACTTTACAG	1140
Db	1119	CGTGGCGTAAAGCCGTAACAGCCAGTGAATTCAGGTTTCGGTGGCTTCAGACTTTACAG	1178
QY	1141	CAGATGCGTGTAAAGCCGCGCTGTCAATTAATACGCCGCCGCGAGAGTGAACAGTACC	1200
Db	1179	CAGATGCGTGTAAAGCCGCGCTGTCAATTAATACGCCGCCGCGAGAGTGAACAGTACC	1238
QY	1201	CTGGCAGAGTGTGAAGAGCCGAATGCCAGGCGCATGTGTTGTTGCCAGGTTTACGCA	1260
Db	1239	CTGGCAGAGTGTGAAGAGCCGAATGCCAGGCGCATGTGTTGTTGCCAGGTTTACGCA	1298
QY	1261	ATCGTAATGAAGCAGCATACCGCGCGGAGGTTTG	1296
Db	1299	ATCGTAATGAAGCAGCATACCGCGCGGAGGTTTG	1334
RESULT 13			
AA28216	AAA28216 standard; DNA; 1489 BP.		
AA28216	12-FEB-2001 (first entry)		
DE	E. coli acidic phosphatase appA2 encoding nucleotide sequence.		
KW	Phosphatase; phytase activity; protease; acidic phosphatase; appA2;		
KM	periplasmic phosphoanhydride phosphohydrolase; animal feed; ds.		
OS	Escherichia coli.		
FT	Key	Location/Qualifiers	
FT	protein_bind	Complement (1..22)	
FT	/*tag= a		
FT	/bound_moiety= "Primer Pfl"		
FT	CDS	16..108	
FT	/*tag= b		
FT	/product= "AppA2 related peptide"		
FT	CDS	184..1482	
FT	/*tag= c		
FT	/product= "AppA2"		
FT	/note= "Acidic phosphatase"		
FT	primer_bind	241..261	
FT	/*tag= d		
FT	/bound_moiety= "Primer E2"		
FT	primer_bind	1471..1489	
FT	/*tag= e		
FT	/bound_moiety= "Primer K2"		
XX	WO200058481-A2.		
XX	05-OCT-2000.		
XX	31-MAR-2000; 2000WO-US08590.		
XX	31-MAR-1999; 99US-0127032.		
XX	(CORR) CORNELL RES FOUND INC.		
XX	Lei X;		
XX	WPI: 2000-619081/59.		
XX	P-PSDB: MAY94753, MAY94756.		
XX	Novel phosphatase having improved phytase activity useful as animal		
XX	feed for improving the accessibility of phosphate to an animal is		
XX	produced by treating phosphatase with protease		
XX	Disclosure; Fig 6; 48pp; English.		
XX	This invention relates to a phosphatase fragment with improved phytase		

CC activity. The phosphatase fragment is created by treating the phosphatase
CC with a protease. The invention includes an *Escherichia coli* nucleotide
CC sequence which encodes an acidic phosphatase appa2 which has improved
CC phosphatase activity. The appa gene of *Escherichia coli* was originally
CC defined as *E. coli* periplasmic phosphohydrolase, phosphatase.
CC although it was not previously known to have phytase activity. The
CC phosphatase fragment can be used in animal feed, to improve the
CC accessibility of phosphate to the animal. The phosphatase gene which
CC encodes the protein can be used in the production of phytase. The present
CC sequence represents the appa2 coding sequence.

XX Sequence 1489 BP; 362 A; 384 C; 403 G; 340 T; 0 other:

Query Match 95.5%; Score 1263.8; DB 21; Length 1489;

Best Local Similarity 97.9%; Pred. No. 0;

Matches 1280; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

```

QY 1 ATGAACGATCTTATCCCATTTTATCTTCTGATCCGTTAACCCGCAATGCA 60
DB 182 ATGAACGATCTTATCCCATTTTATCTTCTGATCCGTTAACCCGCAATGCA 241
QY 61 TTGCTCAGATGAGCGGAGCTGAGCTGGAAGTGTGATTCAGTCATGCT 120
DB 242 TTGCTCAGATGAGCGGAGCTGAGCTGGAAGTGTGATTCAGTCATGCT 301
QY 121 GTGCGTCTCCAAACCAAGGCCACCACTGATGAGATGTACCCGAGACGATGGCA 180
DB 302 GTGCGTCTCCAAACCAAGGCCACCACTGATGAGATGTACCCGAGACGATGGCA 361
QY 181 ACCGCGCGGTTAAACCTGGGTGTGATGACACCGGNGTGTGATGATTCCTATCTC 240
DB 362 ACCGCGCGGTTAAACCTGGGTGTGATGACACCGGNGTGTGATGATTCCTATCTC 421
QY 241 GGAATATTACCAACGCCAGCTGTGATGACGAGATGCTGGGAAAAAGGCTGCCG 300
DB 422 GGAATATTACCAACGCCAGCTGTGATGACGAGATGCTGGGAAAAAGGCTGCCG 481
QY 301 CAGTGTGATGAGCTGCGGATTTATGATGTGACAGACGCTACCCGTAACAGGCGAA 360
DB 482 CAGTGTGATGAGCTGCGGATTTATGATGTGACAGACGCTACCCGTAACAGGCGAA 541
QY 361 GCCGTGCGCGCGGCGGCGGACCTGACTGTGATTAACCCGATACCGAGGATACG 420
DB 542 GCCGTGCGCGCGGCGGCGGACCTGACTGTGATTAACCCGATACCGAGGATACG 601
QY 421 TCCAGTCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCACTGGATACGCG 480
DB 602 TCCAGTCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCACTGGATACGCG 661
QY 481 AACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTTACCGGCAT 540
DB 662 AACGTGACTGACGCGATCTCAGCAGGCGAGGAGGTCATTTGCTGACTTTACCGGCAT 721
QY 541 CGGCAAAACGGCGTTTCCGCAACTGGAACGGGTGTTTATTTTCCGCAATCAATTTGTC 600
DB 722 CGGCAAAACGGCGTTTCCGCAACTGGAACGGGTGTTTATTTTCCCAATCAATTTTGTGC 781
QY 601 CTTAAACGTGAGAAACAGAGAAAGCTGTTTATTACGCGAGGATTAACATCGGAATC 660
DB 782 CTTAAACGTGAGAAACAGAGAAAGCTGTTTATTACGCGAGGATTAACATCGGAATC 841
QY 661 AAGGTGAGCGCGCAATGTTCTATTAAACGGGTGTAAGCCCTGCAATCATGCTGACG 720
DB 842 AAGGTGAGCGCGCAATGTTCTATTAAACGGGTGTAAGCCCTGCAATCATGCTGACG 901
QY 721 GAGATATTCTCTCGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGGAAGATCAAC 780
DB 902 GAAATATTCTCTCGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGGAAGATCACT 961
QY 781 GATTACACAGCAGTGAACACCTTGAAGTTTGCATTAACGCGCAATTTTATTACTACAA 840
DB 962 GATTACACAGCAGTGAACACCTTGAAGTTTGCATTAACGCGCAATTTTATTACTACAA 1021

```

```

QY 841 CGACGCCAGAGGTTGCCCGCAGCCGCCACCCGTTATTGATTGATTGANTCAGCAGC 900
DB 1022 CGCAGCGCAGAGGTTGCCCGCAGCGCCACCCGTTATTGATTGATTGANTCAGCAGC 1081
QY 901 TTGACGCCCGCATTCACCGCAAAACAGCGGATGTGTGATTAACCACTTCAGTCTG 960
DB 1082 TTGACGCCCGCATTCACCGCAAAACAGCGGATGTGTGATTAACCACTTCAGTCTG 1141
QY 961 TTTATTCGCCGACAGTACTAATCTGCAAAATCTGGCGCGCGCATGAGCTCACTGAG 1020
DB 1142 TTTATTCGCCGACAGTACTAATCTGCAAAATCTGGCGCGCGCATGAGCTCACTGAG 1201
QY 1021 ACCTTCCCGTACAGCGGATTAACACCGCCGAGGTGTGATGATGTTTGAACGCTGG 1080
DB 1202 ACCTTCCCGTACAGCGGATTAACACCGCCGAGGTGTGATGATGTTTGAACGCTGG 1261
QY 1081 CGTGGCGTTAAACGATTAACAGCAGTCAAGTTCAGTTCGCGTTCACACTTTCAG 1140
DB 1262 CGTGGCGTTAAACGATTAACAGCAGTCAAGTTCAGTTCGCGTTCACACTTTCAG 1321
QY 1141 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCCCGCGAGAGTGAACCTGACC 1200
DB 1322 CAGATGCGTGAATAAAGCGCGCTGTCAATTAATACGCCCGCGAGAGTGAACCTGACC 1381
QY 1201 CTGGCAGAGTGAAGAGCGAATGCGCAGGCGATGTGTGTTGGCAGGTTTAAACGCA 1260
DB 1382 CTGGCAGAGTGAAGAGCGAATGCGCAGGCGATGTGTGTTGGCAGGTTTAAACGCA 1441
QY 1261 ATCGTGAATGAAGCAGCATCCGCGGTGCTGATGATTCATTCATC 1308
DB 1442 ATCGTGAATGAAGCAGCATCCGCGGTGCTGATGATTCATTCATC 1489

```

RESULT 14

AD06831 standard; DNA; 1486 BP.

AD06831;

06-AUG-2001 (first entry)

DE *E. coli* acid phosphatase/phytase (appa) DNA.

RW Acid phosphatase-phytase; appa; enzymatic activity; food additive;

KW animal feed; monoester phosphate; cereal food; human food; ds.

XX *Escherichia coli*.

XX Key. Location/Qualifiers

FT CDS 16..108

FT /tag- a "Peptide encoded by shorter ORF of appa DNA"

FT CDS 188..1486

FT /tag- b "product- 'E. coli appa protein'"

FT /transl_except- (pos:512..514, aa:Ala)

FT /transl_except- (pos:629..631, aa:Asn)

FT primer_bind complement (245..267)

FT /tag- c "Forward PCR primer E2"

FT primer_bind 568..598

FT /tag- d "Reverse PCR primer A1"

FT primer_bind complement (777..799)

FT /tag- e "Forward PCR primer P2"

FT primer_bind 800..831

FT /tag- f "Reverse PCR primer P3"

FT primer_bind 1473..1486

FT /tag- g "Reverse PCR primer K2"

XX /bound_molety- "Reverse PCR primer K2"

PN W0200136607-A1.
 XX 25-MAY-2001.
 XX 17-NOV-2000; 2000MO-US31622.
 XX 18-NOV-1999; 99US-0166179.
 XX (CORR) CORNELL RES FOUND INC.
 XX Lei X;
 XX WPI; 2001-367572/38.
 DR P-PSDB: AAE02631, AAE02633.
 XX
 PT Muted acid phosphatase/phytase from *Escherichia coli* has improved
 PT enzymatic activity compared to the wild type and is useful as a food
 PT additive, particularly for animal feeds
 XX
 PS Claim 10; Page 7-9; 56pp; English.
 CC The present sequence is a DNA encoding wild type acid phosphatase/phytase
 CC (appa) from *Escherichia coli*. The wild type appa is mutated to enhance
 CC its enzymatic activity. Phytases, a specific group of monoester from
 CC phosphates, are required to initiate the release of phosphate from
 CC phytate, the major storage of phosphate in cereal foods or feeds. The
 CC mutant acid phosphatase/phytase is added to animal feed as a food
 CC additive to improve uptake of phosphate and zinc from the diet. The
 CC phytase can also be added to human food. The mutant phytase is more heat
 CC stable than phytase isolated from *Aspergillus niger* and safer for use
 CC in human food manufacture.
 CC
 SO Sequence 1486 BP; 367 A; 380 C; 403 G; 336 T; 0 other;
 Query Match 95.5%; Score 1263; DB 22; Length 1486;
 Best Local Similarity 98.4%; Pred. No. 0;
 Matches 1275; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 541 CGGCAACGCGCGTTCCGCACTGAAACGGGTCCTTAATTTCCGCAATCAAACTGTGC 600
 DB 728 CGGCAACGCGCGTTCCGCACTGAAACGGGTCCTTAATTTCCGCAATCAAACTGTGC 787
 QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTGATTAACGACGATTTACCATCGAACTC 660
 DB 788 CTTAAACGTGAGAAACAGACGAAAGCTGTGATTAACGACGATTTACCATCGAACTC 847
 QY 661 AAGTGAGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 DB 848 AAGTGAGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
 QY 721 GAGATTTCTTCCTCGCAACAGACGAGGAAATGCGGACCGGGGGGGGAAAGATCAAC 780
 DB 908 GAAATATTTCTTCCTCGCAACAGACGAGGAAATGCGGACCGGGGGGGGAAAGATCAAC 967
 QY 781 GATTCACACAGTGAACACCTGTCATTAAGTTGATTAACGCGCAATTTATTTGCTACAA 840
 DB 968 GATTCACACAGTGAACACCTGTCATTAAGTTGATTAACGCGCAATTTATTTACTACAA 1027
 QY 841 CGCAGCGCAGAGGTTCCCGCAGCGCCGACCCGTTATTTGATTTGATGCGACGC 900
 DB 1028 CGCAGCGCAGAGGTTCCCGCAGCGCCGACCCGTTATTTGATTTGATGCGACGC 1087
 QY 901 TTGAGCGCGCATCCACCGCAAAACAGGGGTATGTTGATTAACCTTCCACTTCAGTACG 960
 DB 1088 TTGAGCGCGCATCCACCGCAAAACAGGGGTATGTTGATTAACCTTCCACTTCAGTACG 1147
 QY 961 TTTATTTCCCGGACACAGATTAATCTGGAATAATCTGCGGCGGCACTGAGGCTCAACTG 1020
 DB 1148 TTTATTTCCCGGACACAGATTAATCTGGAATAATCTGCGGCGGCACTGAGGCTCAACTG 1207
 QY 1021 AGCGTTCCCGGTCAGCGGATTAACAGCGCCGACAGTGTGTAAGTGTGTTGAACGCTG 1080
 DB 1208 AGCGTTCCCGGTCAGCGGATTAACAGCGCCGACAGTGTGTAAGTGTGTTGAACGCTG 1267
 QY 1081 CGTCGCGTAAAGCATTAACAGCGGATTAACAGGTTGCTGCTGCTTCCAGACTTACAG 1140
 DB 1268 CGTCGCGTAAAGCATTAACAGCGGATTAACAGGTTGCTGCTGCTTCCAGACTTACAG 1327
 QY 1141 CAGATCGGTATTAACAGCGCGCTGTATTAATATACCGCCGAGAGGTGAATGAC 1200
 DB 1328 CAGATCGGTATTAACAGCGCGCTGTATTAATATACCGCCGAGAGGTGAATGAC 1387
 QY 1201 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGCGTGGCAGTTTACGCA 1260
 DB 1388 CTGGCAGATGTGAAGAGGAAATGCGCAGGCGATGTGTTGCGTGGCAGTTTACGCA 1447
 QY 1261 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTG 1296
 DB 1448 ATCGTGAATGAAGCAGCATACCGGCGTCAAGTTG 1483

RESULT 15
 AAD06832 standard; DNA; 1486 BP.
 ID AAD06832:
 AC AAD06832:
 XX 06-AUG-2001 (first entry)
 DE E. coli acid phosphatase/phytase (appa) mutant DNA.
 KW Acid phosphatase-phytase; appa; enzymatic activity; food additive;
 OS animal feed; monoester phosphate; cereal food; human food; mutant; ds.
 XX *Escherichia coli*.
 FT Synthetic.
 FT Key Location/Qualifiers
 FT CDS 16..108
 FT /tag- a

Search completed: June 12, 2003, 10:33:46
Job time : 373 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 08:17:53 ; Search time 235 Seconds
(without alignments)
7879.096 Million cell updates/sec

Title: US-09-777-566a-1
Perfect score: 1323
Sequence: 1 atgaagcagctatcctatccc.....atccatcattacattactaa 1323

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 870385 segs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCF_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCFUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Query Result					Description	
No.	Score	Match	Length	ID		
1	1322	99.9	1323	9	US-10-034-985-1	Sequence 1, Appl1
2	1322	99.9	1323	10	US-09-777-566A-1	Sequence 1, Appl1
3	1322	99.9	1323	10	US-09-866-379-1	Sequence 1, Appl1
4	1287.8	97.3	1901	10	US-09-866-379-7	Sequence 7, Appl1
5	1286.2	97.2	1901	10	US-09-866-379-5	Sequence 5, Appl1
6	1286.2	97.2	1901	10	US-09-866-379-9	Sequence 9, Appl1
7	1283	97.0	1901	10	US-09-866-379-6	Sequence 6, Appl1
8	1263.8	95.5	1489	9	US-10-266-041-9	Sequence 9, Appl1
9	218.8	16.5	1325	9	US-10-021-723a-3	Sequence 3, Appl1
10	215.6	16.3	1325	9	US-10-021-723a-1	Sequence 1, Appl1
11	206.6	15.6	1326	9	US-10-021-723a-11	Sequence 11, Appl1
12	159.6	12.1	1431	9	US-10-021-723a-9	Sequence 9, Appl1
13	73	5.5	1230	9	US-10-021-723a-7	Sequence 7, Appl1
14	64.8	4.9	1266	9	US-10-021-723a-5	Sequence 5, Appl1
15	36.4	2.8	802	9	US-10-184-644-312	Sequence 312, App
16	36.4	2.8	802	9	US-10-184-644-312	Sequence 312, App
17	35.4	2.7	2781	9	US-10-092-154-1364	Sequence 1364, App
18	35.4	2.7	2781	10	US-09-764-847-1364	Sequence 1364, App
19	2.6	867	9	US-10-123-155-20	Sequence 20, Appl1	

c	20	33.8	2.6	54	9	US-10-021-723a-47	Sequence 47, Appl
	21	33.8	2.6	450	9	US-09-796-692-7902	Sequence 7902, Ap
	22	33.8	2.6	450	9	US-10-040-862-7902	Sequence 7902, Ap
	23	33.8	2.6	508	9	US-09-796-692-8272	Sequence 8272, Ap
	24	33.8	2.6	508	9	US-10-040-862-8272	Sequence 8272, Ap
	25	33.4	2.5	974	10	US-09-770-445-281	Sequence 281, Appl
	26	33	2.5	790	9	US-10-123-155-204	Sequence 204, Appl
	27	32.6	2.5	55	9	US-10-021-723a-22	Sequence 22, Appl
	28	32.6	2.5	14800	10	US-09-954-456-1601	Sequence 1601, Ap
	29	32.4	2.4	464	9	US-09-918-995-12838	Sequence 12838, A
	30	32.4	2.4	607	9	US-10-123-155-344	Sequence 344, Appl
	31	32.4	2.4	888	9	US-10-123-155-544	Sequence 544, Appl
	32	32.2	2.4	162	9	US-09-828-523a-97	Sequence 97, Appl
	33	32.2	2.4	162	9	US-09-966-521-89	Sequence 89, Appl
	34	32.2	2.4	401	10	US-09-969-708-272	Sequence 272, Appl
	35	32.2	2.4	401	10	US-09-880-107-1305	Sequence 1305, Ap
	36	32.2	2.4	457	9	US-10-123-155-236	Sequence 236, Appl
	37	31.8	2.4	614	10	US-09-974-300-2058	Sequence 2058, Ap
	38	31.8	2.4	1049	9	US-10-123-155-358	Sequence 358, Appl
	39	31.8	2.4	7558	9	US-10-091-458-52	Sequence 52, Appl
	40	31.6	2.4	3209	10	US-09-925-301-474	Sequence 474, Appl
	41	31.6	2.4	8460	9	US-10-237-271-2	Sequence 2, Appl1
	42	31.4	2.4	625	9	US-10-123-155-70	Sequence 70, Appl
	43	31.4	2.4	2890	10	US-09-881-752a-361	Sequence 361, Appl
	44	31.2	2.4	1509	9	US-09-828-523a-89	Sequence 89, Appl
	45	31.2	2.4	1509	9	US-09-966-521-83	Sequence 83, Appl

ALIGNMENTS

RESULT 1	US-10-034-985-1
Sequence 1, Application US/10034985	
Publication No. US2003004981SAI	
GENERAL INFORMATION:	
APPLICANT: Kretz, Keith	
TITLE OF INVENTION: NOVEL PHRYTASE	
FILE REFERENCE: 09010/029003	
CURRENT APPLICATION NUMBER: US/10/034,985	
CURRENT FILING DATE: 2001-12-21	
PRIOR APPLICATION NUMBER: US/09/580,515	
PRIOR FILING DATE: 1999-05-25	
PRIOR APPLICATION NUMBER: 09/291,931	
PRIOR FILING DATE: 1999-04-13	
PRIOR APPLICATION NUMBER: 08/910,798	
PRIOR FILING DATE: 1997-08-13	
PRIOR APPLICATION NUMBER: 09/259,214	
PRIOR FILING DATE: 1999-03-01	
NUMBER OF SEQ ID NOS: 4	
SOFTWARE: FastSeq for Windows Version 4.0	
SEQ ID NO 1	
LENGTH: 1323	
TYPE: DNA	
ORGANISM: Escherichia coli	
FEATURE:	
NAME/KEY: CDS	
LOCATION: (1)..(1320)	
NAME/KEY: misc_feature	
LOCATION: (1)..(1323)	
OTHER INFORMATION: n = A,T,C or G	
US-10-034-985-1	
Query Match	99.9%; Score 1322; DB 9; Length 1323;
Best local Similarity	100.0%; Pred. No. 0;
Matches 1323; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGTTAACCGGCAATCGCA 60
Db	1 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGTTAACCGGCAATCGCA 60
Oy	61 TTCGCTAGAGTGGAGCGGAGCTGAAGTGTGATTCAGTCGTCATGCT 120

61 TTCGCTCAGATGACCGGAGCTGAGCTGGAAGTGTGATGTCAGTCGTCAGT 120
 QY 121 GTGGGTGCTCCACCAAGGCGCCCACTGATGAGATGTACCCAGACGATGCGCA 180
 Db 121 GTGGGTGCTCCACCAAGGCGCCCACTGATGAGATGTACCCAGACGATGCGCA 180
 QY 181 ACCGTGGCGGTAAACTGGGTGGCTGACACCGGNGGTGTGAGCTTAATGCGCTATCTC 240
 Db 181 ACCGTGGCGGTAAACTGGGTGGCTGACACCGGNGGTGTGAGCTTAATGCGCTATCTC 240
 QY 241 GGACATTACCAAGCGCAGCTGTGTAGCCAGCGATGCTGGGAAAAGGGCTGCCCG 300
 Db 241 GGACATTACCAAGCGCAGCTGTGTAGCCAGCGATGCTGGGAAAAGGGCTGCCCG 300
 QY 301 CAGTCTGTGTCAGTGGCGATTAATGCTGATGTGACGAGCGCTACCCGTAACAGCGCA 360
 Db 301 CAGTCTGTGTCAGTGGCGATTAATGCTGATGTGACGAGCGCTACCCGTAACAGCGCA 360
 QY 361 GCGTTCGCGCGCGGCTGGACCTGACCTGACATGACCAATACCGTACATACCGAGCATACG 420
 Db 361 GCGTTCGCGCGCGGCTGGACCTGACCTGACATGACCAATACCGTACATACCGAGCATACG 420
 QY 421 TCCAGTCCCGATCCGTTATTAATCTCTAATAAACTGCGTTGGCACTGGATTAACCG 480
 Db 421 TCCAGTCCCGATCCGTTATTAATCTCTAATAAACTGCGTTGGCACTGGATTAACCG 480
 QY 481 AACGTGATGACGCGATCTCAGCAGGAGGAGGATGATGCTGATTAACCGGCGAT 540
 Db 481 AACGTGATGACGCGATCTCAGCAGGAGGAGGATGATGCTGATTAACCGGCGAT 540
 QY 541 CCGGAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
 Db 541 CCGGAAACGGCGTTTCGGGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
 QY 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACGAGGAACTC 660
 Db 601 CTTAAGCGTGAAGAAACAGAGGAAAGCTGTTCAATTAACGAGGATTAACGAGGAACTC 660
 QY 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGGGATGAGCTGATGATGATGATGATG 720
 Db 661 AAGGTGAGCGCGCAATGCTCATTAACCGGTGGGATGAGCTGATGATGATGATGATG 720
 QY 721 GAGATATTCTCTGCAACAGACAGGGAATGCGGAGCGGGGTGGGGAAGATCAAC 780
 Db 721 GAGATATTCTCTGCAACAGACAGGGAATGCGGAGCGGGGTGGGGAAGATCAAC 780
 QY 781 GATTCACACACAGTGAACACCTTGTGATGATTAACGAGGATTAATTTGCTACAA 840
 Db 781 GATTCACACACAGTGAACACCTTGTGATGATTAACGAGGATTAATTTGCTACAA 840
 QY 841 CGCAGCGCGAGAGTGGCCCGCAGCGCCACCCGTTATGATGATGATGATGATGATG 900
 Db 841 CGCAGCGCGAGAGTGGCCCGCAGCGCCACCCGTTATGATGATGATGATGATGATG 900
 QY 901 TTGACGCGCCATCCACCGCAAAACAGCGGATGATGATGATGATGATGATGATG 960
 Db 901 TTGACGCGCCATCCACCGCAAAACAGCGGATGATGATGATGATGATGATGATG 960
 QY 961 TTTATGCGCGGACAGTACTAATCTGCAAAATCTGCGGCGCGACAGTGAAGCTG 1020
 Db 961 TTTATGCGCGGACAGTACTAATCTGCAAAATCTGCGGCGCGACAGTGAAGCTG 1020
 QY 1021 ACCTGTCGCGGTACGCGGATTAACAGCGCGGATGATGATGATGATGATGATG 1080
 Db 1021 ACCTGTCGCGGTACGCGGATTAACAGCGCGGATGATGATGATGATGATGATG 1080
 QY 1081 CGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 Db 1081 CGTGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1140
 QY 1141 CAGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
 Db 1141 CAGATGCGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200

QY 1201 CTGGCAGATGTGAAGACGGAATGCGAGGCGATGTGTTGCTGGCAGGTTTACGCA 1260
 Db 1201 CTGGCAGATGTGAAGACGGAATGCGAGGCGATGTGTTGCTGGCAGGTTTACGCA 1260
 QY 1261 ATCGGATGTAAGCAGCATACCGCGCTGACGTTGATGATGATGATGATGATGATG 1320
 Db 1261 ATCGGATGTAAGCAGCATACCGCGCTGACGTTGATGATGATGATGATGATGATG 1320
 QY 1321 TAA 1323
 Db 1321 TAA 1323
 RESULT 2
 US-09-777-566a-1
 ; Sequence 1, Application US/0977566A
 ; Patent No. US20010055788A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: KREZ, Jay
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHAGES AND USES THEREOF
 ; FILE REFERENCE: DIVER370-6
 ; CURRENT FILING DATE: US/09/777,566A
 ; PRIOR FILING DATE: 2001-06-11
 ; PRIOR FILING DATE: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR FILING DATE: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 1323
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1320)
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1323)
 ; OTHER INFORMATION: n is any nucleotide
 US-09-777-566a-1
 Query Match 99.9%; Score 1322; DB 10; Length 1323;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGAAGGAGTCTTAATCCCAATTTTATCTCTGATGATGATGATGATGATGATG 60
 Db 1 ATGAAGGAGTCTTAATCCCAATTTTATCTCTGATGATGATGATGATGATGATG 60
 QY 61 TTCGCTCAGATGACCGGAGCTGAGCTGGAAGTGTGATGATGATGATGATGATG 120
 Db 61 TTCGCTCAGATGACCGGAGCTGAGCTGGAAGTGTGATGATGATGATGATGATG 120
 QY 121 GTGGGTGCTCCACCAAGGCGCCCACTGATGAGATGTACCCAGACGATGCGCA 180
 Db 121 GTGGGTGCTCCACCAAGGCGCCCACTGATGAGATGTACCCAGACGATGCGCA 180
 QY 181 ACCTGTCGCGGTAAACTGGGTGGCTGACACCGGNGGTGTGAGCTTAATGCGCTATCTC 240
 Db 181 ACCTGTCGCGGTAAACTGGGTGGCTGACACCGGNGGTGTGAGCTTAATGCGCTATCTC 240
 QY 241 GGACATTACCAAGCGCAGCTGTGTAGCCAGCGATGCTGGGAAAAGGGCTGCCCG 300
 Db 241 GGACATTACCAAGCGCAGCTGTGTAGCCAGCGATGCTGGGAAAAGGGCTGCCCG 300
 QY 301 CAGTCTGTGTCAGTGGCGATTAATGCTGATGTGACGAGCGCTACCCGTAACAGCGCA 360

|||||
Db 301 CAGTCTGTCAGTCTGCGATTAATATGCTGATGTGAGAGAGCTACCCGTAATAACAGCGAA 360
OY 361 GCCTTGCCGCGCGGGCTGGCAGCTGACTGCAATACCGTACATACCGAGAGATAGC 420
Db 361 GCCTTGCCGCGCGGGCTGGCAGCTGACTGCAATACCGTACATACCGAGAGATAGC 420
OY 421 TCCAGTCCGAGTCCGATTAATATGCTGATGTGAGAGAGCTACCCGTAATAACAGCG 480
Db 421 TCCAGTCCGAGTCCGATTAATATGCTGATGTGAGAGAGCTACCCGTAATAACAGCG 480
OY 481 AACGTGACTGACGCGATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 481 AACGTGACTGACGCGATCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
OY 541 CGGCAAAAGCGCGCTTGCGCAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 541 CGGCAAAAGCGCGCTTGCGCAACTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
OY 601 CTTAAAGTGAAGAAACAG 660
Db 601 CTTAAAGTGAAGAAACAG 660
OY 661 AAGGTGAGCGCGCAGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
Db 661 AAGGTGAGCGCGCAGCAATGTCTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
OY 721 GAGATTTTCTCTGTCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 721 GAGATTTTCTCTGTCGCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
OY 781 GATTACACAG 840
Db 781 GATTACACAG 840
OY 841 CGCAGCCAG 900
Db 841 CGCAGCCAG 900
OY 901 TTGACGCCCGCTACACAG 960
Db 901 TTGACGCCCGCTACACAG 960
OY 961 TTGATTTGCGGAG 1020
Db 961 TTGATTTGCGGAG 1020
OY 1021 ACGCTTCCGCGTACAG 1080
Db 1021 ACGCTTCCGCGTACAG 1080
OY 1081 CCGCGGCTAAGCAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
Db 1081 CCGCGGCTAAGCAGATTAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
OY 1141 CAGATCGGTGATTAACAG 1200
Db 1141 CAGATCGGTGATTAACAG 1200
OY 1201 CTTGCGAGAGATTAACAG 1260
Db 1201 CTTGCGAGAGATTAACAG 1260
OY 1261 ATGCTGATTAACAG 1320
Db 1261 ATGCTGATTAACAG 1320
OY 1321 TAA 1323
Db 1321 TAA 1323

US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ. ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1
Query Match 99.9%; Score 1322; DB 10; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGCTTAACCCCGCATCTGCA 60
Db 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGCTTAACCCCGCATCTGCA 60
OY 61 TTGCTCAGAGTAGCGCGAG 120
Db 61 TTGCTCAGAGTAGCGCGAG 120
OY 121 GTGCTGCTCAACCAAGGCGACGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 121 GTGCTGCTCAACCAAGGCGACGCACTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
OY 181 ACCTGCGGCTAAG 240
Db 181 ACCTGCGGCTAAG 240
OY 241 GGAACATTAACCAAGCGCGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 241 GGAACATTAACCAAGCGCGCTGTGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
OY 301 CAGTCTGCTAGAGTCCGAGATTAATGCTGATGTCGAGAGAGAGAGAGAGAGAGAGAGAG 360
Db 301 CAGTCTGCTAGAGTCCGAGATTAATGCTGATGTCGAGAGAGAGAGAGAGAGAGAGAGAG 360
OY 361 GCCTTGCCGCGCGGGCTGGCAGCTGACTGCAATACCGTACATACCGAGAGATAGC 420
Db 361 GCCTTGCCGCGCGGGCTGGCAGCTGACTGCAATACCGTACATACCGAGAGATAGC 420
OY 421 TCCAGTCCGAGTCCGATTAATATGCTGATGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480

```

Db      421  TCCAGTCCCATCCGTTATTTAAATCCCTAAAAAAGCGCTTTGCCAAGCGATTAACGCG 480
OY      481  AACGTGACGAGCGATCCTCAGACGGGAGGAGGCGATTTGCTGTACCGGGCAT 540
Db      481  AACGTGACGAGCGATCCTCAGACGGGAGGAGGCGATTTGCTGTACCGGGCAT 540
OY      541  CGGCAAAAGCGGCTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
Db      541  CGGCAAAAGCGGCTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
OY      601  CTTAAACGTGAGAAACAGGAGCAAAAGCTGTTCAATTAACGAGCATTAACATGGAATCTC 660
Db      601  CTTAAACGTGAGAAACAGGAGCAAAAGCTGTTCAATTAACGAGCATTAACATGGAATCTC 660
OY      661  AAGGTGAGCGCCCAATATGCTCATTTAACGGGTGCGGTAAAGCTCGCATCAATGCTACG 720
Db      661  AAGGTGAGCGCCCAATATGCTCATTTAACGGGTGCGGTAAAGCTCGCATCAATGCTACG 720
OY      721  GAGATATTTCTCTGCAACAGCAGCAGGAATGCGGAGCGGGGTGGGAAGGATCACC 780
Db      721  GAGATATTTCTCTGCAACAGCAGCAGGAATGCGGAGCGGGGTGGGAAGGATCACC 780
OY      781  GATTACACCAAGTGAACACACTTGTCTAAGTTTGATTAACGCGCAATTTATTGCTACAA 840
Db      781  GATTACACCAAGTGAACACACTTGTCTAAGTTTGATTAACGCGCAATTTATTGCTACAA 840
OY      841  CGCAGCGCAGAGGTGGCCCGCAGCGCGCCGACCCCGTTATTTGATTTGATCATGCGACG 900
Db      841  CGCAGCGCAGAGGTGGCCCGCAGCGCGCCGACCCCGTTATTTGATTTGATCATGCGACG 900
OY      901  TTGACGCCCATCCACCGCAAAAACAGCGCTATGCTGACATTAACCCACTTCAGTACTG 960
Db      901  TTGACGCCCATCCACCGCAAAAACAGCGCTATGCTGACATTAACCCACTTCAGTACTG 960
OY      961  TTTATTTGCGGAGACGATTAATCTGCGCAAAATCTGCGCGCGCGACGAGGAGCTCACTG 1020
Db      961  TTTATTTGCGGAGACGATTAATCTGCGCAAAATCTGCGCGCGCGACGAGGAGCTCACTG 1020
OY      1021  ACGCTTCCCGGTACGCGGATTAACACGCGCCGACAGGTGATGAATGCTGTTGAACGCTG 1080
Db      1021  ACGCTTCCCGGTACGCGGATTAACACGCGCGCCGACAGGTGATGAATGCTGTTGAACGCTG 1080
OY      1081  CGTGGCGTAAAGCATTAACAGCAGTGTGATTCAGGTTCTGCTGCTTCAGACTTTTACAG 1140
Db      1081  CGTGGCGTAAAGCATTAACAGCAGTGTGATTCAGGTTCTGCTGCTTCAGACTTTTACAG 1140
OY      1141  CAGATGCGTATTAACAGCGCGCTGTCTATTAATTAACGCGCGCGCGAGAGGTGAACCTGAC 1200
Db      1141  CAGATGCGTATTAACAGCGCGCTGTCTATTAATTAACGCGCGCGCGAGAGGTGAACCTGAC 1200
OY      1201  CTGGCAGAGATGTGAAGAGCAAAATGCGCAGGGCATGTGTCGTTGCGAGGTTTTACGCA 1260
Db      1201  CTGGCAGAGATGTGAAGAGCAAAATGCGCAGGGCATGTGTCGTTGCGAGGTTTTACGCA 1260
OY      1261  ATGTGTAATTAACAGCAGCATACCGGCGCTGCGAGTTGAGATCTCATCAACATCAAC 1320
Db      1261  ATGTGTAATTAACAGCAGCATACCGGCGCTGCGAGTTGAGATCTCATCAACATCAAC 1320
OY      1321  TAA 1323
Db      1321  TAA 1323

```

RESULT 4

```

US-09-866-379-7
; Sequence 7, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin

```

```

; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-7

Query Match      97.3%; Score 1287.8; DB 10; Length 1901;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 1293; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY      1  ATGAACGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCGCAATCTGCA 60
Db      188  ATGAACGATCTTAATCCATTTTATCTTCTGATTCGTTAAACCGCAATCTGCA 247
OY      61  TTGCTCAGAGTGAGCCGAGCTGAGCTGGAAGTGTGATTTGATGCTGATGCT 120
Db      248  TTGCTCAGAGTGAGCCGAGCTGAGCTGGAAGTGTGATTTGATGCTGATGCT 307
OY      121  GTGCTGCTCCAAACAGGCGCAGCAGCACTGATGAGATGTCACCCAGAGCATG 180
Db      248  GTGCTGCTCCAAACAGGCGCAGCAGCACTGATGAGATGTCACCCAGAGCATG 307
OY      308  GTGCTGCTCCAAACAGGCGCAGCAGCACTGATGAGATGTCACCCAGAGCATG 367
Db      181  ACCTGCGCGGTAAACTGGGTGCTGACACCGCGNGTGGTGAAGTAATCGCTATCTC 240
OY      368  ACCTGCGCGGTAAACTGGGTGCTGACACCGCGNGTGGTGAAGTAATCGCTATCTC 427
Db      241  GGACATTACCAAGCCGAGCTGTGATGCGGAGGATTTGCTGGGGAAGAGGCTGCG 300
OY      428  GGACATTACCAAGCCGAGCTGTGATGCGGAGGATTTGCTGGGGAAGAGGCTGCG 487
Db      301  CAGTCTGTCAGAGCGGATTAATGCTGATGTGACGAGGATACCCGTTAAACAGG 360
OY      488  CAGTCTGTCAGAGCGGATTAATGCTGATGTGACGAGGATACCCGTTAAACAGG 547
Db      361  GCCTTCCGCGCGGCTGCGACCTGACTGTGCAATTAACGTTACACGAGATACG 420
OY      548  GCCTTCCGCGCGGCTGCGACCTGACTGTGCAATTAACGTTACACGAGATACG 607
Db      421  TCCAGTCCCATCCGTTATTTAAATCCCTAAAAAAGCGCTTTGCCAAGCGATTA 480
OY      608  TCCAGTCCCATCCGTTATTTAAATCCCTAAAAAAGCGCTTTGCCAAGCGATTA 667
Db      481  AACGTGACGAGCGATCCTCAGACGGGAGGAGGCGATTTGCTGTACCGGGCAT 540
OY      668  AACGTGACGAGCGATCCTCAGACGGGAGGAGGCGATTTGCTGTACCGGGCAT 727
Db      541  CGGCAAAAGCGGCTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACT 600
OY      728  CGGCAAAAGCGGCTTTCCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACT 787

```

QY 601 CTTAAAGCTGAGAAAGAGCAAGAGCTGTTCAATTAACGAGCATTAACATCGAATC 660
 Db 788 CTTAAAGCTGAGAAAGAGCAAGAGCTGTTCAATTAACGAGCATTAACATCGAATC 847
 QY 661 AAGGTAGGCGCGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 Db 848 AAGGTAGGCGCGCGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAGCAAGAGAGGAAATGCGGAGCGGGGTGGGGAAGATACG 780
 Db 908 GAGATATTTCTCTGCAACAGCAAGAGAGGAAATGCGGAGCGGGGTGGGGAAGATACG 967
 QY 781 GATTCAACACAGTGAACACCTTGTCTAATTTGCAATACGCGCAATTTATTTGCTACAA 840
 Db 968 GATTCAACACAGTGAACACCTTGTCTAATTTGCAATACGCGCAATTTATTTGCTACAA 1027
 QY 841 CGACAGCGAGAGTGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900
 Db 1028 CGACAGCGAGAGTGGCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1087
 QY 901 TTGAGCG 960
 Db 1088 TTGAGCG 1147
 QY 961 TTGATGCGCGAGACGATTAATCTGGCAAAATCTCGCGCGCGCGCGCGCGCGCGCGCG 1020
 Db 1148 TTGATGCGCGAGACGATTAATCTGGCAAAATCTCGCGCGCGCGCGCGCGCGCGCGCG 1207
 QY 1021 ACCGTTCCCGGTCAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080
 Db 1208 ACCGTTCCCGGTCAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1267
 QY 1081 CGTCGCGGTAGCGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1140
 Db 1268 CGTCGCGGTAGCGATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCG 1327
 QY 1141 CAGATCGGTATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200
 Db 1328 CAGATCGGTATTAACAGCGCGGATTAACAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1387
 QY 1201 CTGGCAGATGTAAGAGCGCAATGCGCGAGGCGATGTCGTTGGCAGGTTTACGCA 1260
 Db 1388 CTGGCAGATGTAAGAGCGCAATGCGCGAGGCGATGTCGTTGGCAGGTTTACGCA 1447
 QY 1261 ATCGTAATGAACAGCATACCGCGGTGCGAGTTTGAT 1300
 Db 1448 ATCGTAATGAACAGCATACCGCGGTGCGAGTTTGAT 1487

RESULT 5
 US-09-866-379-5
 ; Sequence 5, Application US/09866379
 ; Patent No. US20020136754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KRETZ, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT APPLICATION NUMBER: US/09/866,379
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: US 09/580,515
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTGATTCCTGTTAAACCCCGCAATGCA 60
 Db 188 ATGAAAGCGATCTTAATCCCATTTTATCTCTGATTCCTGTTAAACCCCGCAATGCA 247
 QY 61 TTGCTGAGAGTGAAGCGCGAGCTGAACCTGGAAGTGTGATTTGATGCTGATGCT 120
 Db 248 TTGCTGAGAGTGAAGCGCGAGCTGAACCTGGAAGTGTGATTTGATGCTGATGCT 307
 QY 121 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGTCACCCAGACGATGCGCA 180
 Db 308 GTGCGTCTCCAAACCAAGGCGACGCACTGATGAGATGTCACCCAGACGATGCGCA 367
 QY 181 ACCGCGCGGTTAAACTGAGTGGTGTGACACCGCGGAGTGGTGAAGTAAATGCGCTATCTC 240
 Db 368 ACCGCGCGGTTAAACTGAGTGGTGTGACACCGCGGAGTGGTGAAGTAAATGCGCTATCTC 427
 QY 241 GGACATTAACCAAGCGCGAGCTGCTGTGACGCGAGGATGCTGCGGGAAGAGGCTGCCG 300
 Db 428 GGACATTAACCAAGCGCGAGCTGCTGTGACGCGAGGATGCTGCGGGAAGAGGCTGCCG 487
 QY 301 CAGTCTGCTGAGTGGGATTAATGCTGATGTCAGCAGGAGTACCGGTAACAGGCGCA 360
 Db 488 CAGTCTGCTGAGTGGGATTAATGCTGATGTCAGCAGGAGTACCGGTAACAGGCGCA 547
 QY 361 GCGTTCCCGCGGCGTGGCAGCTGACCTGATGCAATTAACCGTATACCGAGCGATACG 420
 Db 548 GCGTTCCCGCGGCGTGGCAGCTGACCTGATGCAATTAACCGTATACCGAGCGATACG 607
 QY 421 TCAGTCCCGGATGCTGATTAATGCTGTAACAACTGCGGTTGCCAATGATACGCG 480
 Db 608 TCAGTCCCGGATGCTGATTAATGCTGTAACAACTGCGGTTGCCAATGATACGCG 667
 QY 481 AACGTGACTGACGGATCTCAGAGGCGAGGAGGTCATTTGCTGACTTACCGGGCAT 540
 Db 668 AACGTGACTGACGGATCTCAGAGGCGAGGAGGTCATTTGCTGACTTACCGGGCAT 727
 QY 541 CGGCAAAAGCGGTTTCGCAAACTGGAAGCGGTCGTTAATTTTCGCAATCAACTTTCG 600
 Db 728 CGGCAAAAGCGGTTTCGCAAACTGGAAGCGGTCGTTAATTTTCGCAATCAACTTTCG 787
 QY 601 CTTAAAGCTGAGAAAGAGCAAGAAAGCTGTTCAATTAACGAGCATTAACATCGAATC 660
 Db 788 CTTAAAGCTGAGAAAGAGCAAGAAAGCTGTTCAATTAACGAGCATTAACATCGAATC 847
 QY 661 AAGGTAGGCGCGCGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 720
 Db 848 AAGGTAGGCGCGCGACAAATGTCATTAACCGGTGCGGTAAAGCTCGCATCAATGCTGACG 907
 QY 721 GAGATATTTCTCTGCAACAGCAAGAGAGGAAATGCGGAGCGGGGTGGGGAAGATACG 780
 Db 908 GAGATATTTCTCTGCAACAGCAAGAGAGGAAATGCGGAGCGGGGTGGGGAAGATACG 967
 QY 781 GATTCAACACAGTGAACACCTTGTCTAATTTGCAATACGCGCAATTTATTTGCTACAA 840
 Db 968 GATTCAACACAGTGAACACCTTGTCTAATTTGCAATACGCGCAATTTATTTGCTACAA 1027

;; PRIOR APPLICATION NUMBER: US 08/910,798
 ;; PRIOR FILING DATE: 1997-08-13
 ;; NUMBER OF SEQ ID NOS: 10
 ;; SOFTWARE: Patent version 3.1
 ;; SEQ ID NO 5
 ;; LENGTH: 1901
 ;; TYPE: DNA
 ;; ORGANISM: Escherichia coli
 ;; FEATURE:
 ;; NAME/KEY: misc_feature
 ;; LOCATION: (1)..(1901)
 ;; OTHER INFORMATION: n is any nucleotide
 US-09-866-379-5

Query Match 97.2% Score 1286.2; DB 10; Length 1901;
 Best Local Similarity 99.4% Pred. No. 0;
 Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

QY 841 CGCAGCCAGAGTTGGCCCGCAGCCGCCACCCTTATTTGATTTGATCGACGC 900
|||||
1028 CGCAGCCAGAGTTGGCCCGCAGCCGCCACCCTTATTTGATTTGATCGACGC 1087
QY 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGACATTTACCACTTCACTG 960
|||||
1088 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGACATTTACCACTTCACTG 1147
QY 961 TTTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGCACTGAGACTCACTG 1020
|||||
1148 TTTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGCACTGAGACTCACTG 1207
QY 1021 ACCGTTCCCGGTACGCCGATTAACACCGCCGCCAGGTGTGAATCTGTTTGAACGCTG 1080
|||||
1208 ACCGTTCCCGGTACGCCGATTAACACCGCCGCCAGGTGTGAATCTGTTTGAACGCTG 1267
QY 1081 CGTGGGTAAAGCATTAACACCGCATGTGATTCAGTTTCGCTGCTTCGACACTTACAG 1140
|||||
1268 CGTGGGTAAAGCATTAACACCGCATGTGATTCAGTTTCGCTGCTTCGACACTTACAG 1327
QY 1141 CAGATGGGTAAAGCATTAACACCGCATGTGATTCAGTTTCGCTGCTTCGACACTTACAG 1200
|||||
1328 CAGATGGGTAAAGCATTAACACCGCATGTGATTCAGTTTCGCTGCTTCGACACTTACAG 1387
QY 1201 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGGCATGTGCTGTCGTCGACAGTTTACGCAA 1260
|||||
1388 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGGCATGTGCTGTCGTCGACAGTTTACGCAA 1447
QY 1261 ATGCTGAATGAAGCATTAACCGCGCTGAGTTTGAT 1300
|||||
1448 ATGCTGAATGAAGCATTAACCGCGCTGAGTTTGAT 1487

RESULT 6
US-09-866-379-9

Sequence 9, Application US/09866379
Patent No. US20020136754A1

GENERAL INFORMATION:

APPLICANT: DIVERSA CORPORATION
APPLICANT: SHORT, Jay
APPLICANT: KRETZ, Keith
APPLICANT: GRAY, Kevin
APPLICANT: BARTON, Nelson
APPLICANT: GARRETT, James
TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
FILE REFERENCE: DIVER1370-7
CURRENT APPLICATION NUMBER: US/09/866,379
PRIORITY FILING DATE: 2001-05-24
PRIORITY APPLICATION NUMBER: US 09/580,515
PRIORITY FILING DATE: 2000-05-25
PRIORITY APPLICATION NUMBER: US 09/318,528
PRIORITY FILING DATE: 1999-05-25
PRIORITY APPLICATION NUMBER: US 09/291,931
PRIORITY FILING DATE: 1999-04-13
PRIORITY APPLICATION NUMBER: US 09/259,214
PRIORITY FILING DATE: 1999-03-01
PRIORITY APPLICATION NUMBER: US 08/910,798
PRIORITY FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn version 3.1

SEQ ID NO 9
LENGTH: 1901
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)..(1901)
OTHER INFORMATION: n is any nucleotide
US-09-866-379-9

Query Match 97.28; Score 1286.2; DB 10; Length 1901;

Best Local Similarity 99.48; Pred. No. 0; Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 1 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTACCCGCAATCTGCA 60
188 ATGAAGCGATCTTAATCCATTTTATCTTCTGATTCGTTACCCGCAATCTGCA 247
QY 61 TTGCTAGAGTGAGCCGAGCTGACAGCTGGAAGTGTGATTTGATTCGATGCT 120
248 TTGCTAGAGTGAGCCGAGCTGACAGCTGGAAGTGTGATTTGATTCGATGCT 307
QY 121 GTGCTGCTCCAAACCAAGGCGACCACTGATGAGATGCAACCCGACGCTGSCCA 180
308 GTGCTGCTCCAAACCAAGGCGACCACTGATGAGATGCAACCCGACGCTGSCCA 367
QY 181 ACCGTGGCGGTAATCACTGAGTGTGCTGACACCGCGNGTGTGAGTGAATCGCTATCTC 240
368 ACCGTGGCGGTAATCACTGAGTGTGCTGACACCGCGNGTGTGAGTGAATCGCTATCTC 427
QY 241 GGACATTACCAACGCCACGCTGTGTAGCCGAGGATTCGCGGCAAAAAGGCTGCCG 300
428 GGACATTACCAACGCCACGCTGTGTAGCCGAGGATTCGCGGCAAAAAGGCTGCCG 487
QY 301 CAGTCTGCTAGCTGCGGATTAATGCTGATGTGAGAGGCTACCGTAAACAGGCGAA 360
488 CAGTCTGCTAGCTGCGGATTAATGCTGATGTGAGAGGCTACCGTAAACAGGCGAA 547
QY 361 GCCTTTCGCGCGCGGCTGCGACCTGACTGTGCAATAACCGTATACCAAGCAGATACG 420
548 GCCTTTCGCGCGCGGCTGCGACCTGACTGTGCAATAACCGTATACCAAGCAGATACG 607
QY 421 TCCAGTCCCGATCCGTTATTTATCTCTCTTAAAACTGGCTTCCCAACTGATGAACGGC 480
608 TCCAGTCCCGATCCGTTATTTATCTCTCTTAAAACTGGCTTCCCAACTGATGAACGGC 667
QY 481 AAGCTGACTGACGCGATCTCAGCAGGAGGAGGCTCAATGCTGACTTATACCGGCGAT 540
668 AAGCTGACTGACGCGATCTCAGCAGGAGGAGGCTCAATGCTGACTTATACCGGCGAT 727
QY 541 CGGCAAAAGCGGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC 600
728 CGGCAAAAGCGGTTTCGCAACTGGAACGGGTGCTTAATTTCCGCAATCAACTGTGC 787
QY 601 CTTAAAGTGGAAGAACAGAGCAAGAGTGTCAATTAACGACGATTAACATCGGAATCTC 660
788 CTTAAAGTGGAAGAACAGAGCAAGAGTGTCAATTAACGACGATTAACATCGGAATCTC 847
QY 661 AAGGTGAGCGCGCAATGTCTCATTTAACGGTGTGAGGCTTCCGATCAATGCTGACG 720
848 AAGGTGAGCGCGCAATGTCTCATTTAACGGTGTGAGGCTTCCGATCAATGCTGACG 907
QY 721 GAGATATTTCTCTGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGAGGATCAAC 780
908 GAGATATTTCTCTGCAACAAGCAGAGGAATGCGGAGCCGGGTGGGAGGATCAAC 967
QY 781 GATTCACACAGAGTGGAACACCTTCTCAATGATTAACGCGCAATTTTATTTGCTCAAA 840
968 GATTCACACAGAGTGGAACACCTTCTCAATGATTAACGCGCAATTTTATTTGCTCAAA 1027
QY 841 CGCAGCGCAGAGTTGGCCCGCAGCCGCCACCCTTATTTGATTTGATCGACGC 900
1028 CGCAGCGCAGAGTTGGCCCGCAGCCGCCACCCTTATTTGATTTGATCGACGC 1087
QY 901 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGACATTTACCACTTCACTG 960
1088 TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGACATTTACCACTTCACTG 1147
QY 961 TTTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGCACTGAGACTCACTG 1020
1148 TTTATTTCCGGACACGATTAATCTGCAAAATCTGCGCGCACTGAGACTCACTG 1207
QY 1021 ACCGTTCCCGGTACGCCGATTAACACCGCCGCCAGGTGTGAATCTGTTTGAACGCTG 1080

Db 1208 ACCTTCCTCCGCTGACCGGATTAACACGCCGCCAGGTGTGTAAGTGTGTTGAACGCTGG 1267
Qy 1081 CGTCGGCTAAGCATTAACACGAGTGTGATTCAGGTTTCGCTGCTTCACAGACTTTCAG 1140
Db 1268 CGTCGGCTAAGCATTAACACGAGTGTGATTCAGGTTTCGCTGCTTCACAGACTTTCAG 1337
Qy 1141 CAGATGCGGTATTAACACGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1200
Db 1328 CAGATGCGGTATTAACACGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1387
Qy 1201 CTGGCAGAGTGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCA 1260
Db 1388 CTGGCAGAGTGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCA 1447
Qy 1261 ATCGTGAATGAGCAGCATACCGCGCTGAGTTTGAT 1300
Db 1448 ATCGTGAATGAGCAGCATACCGCGCTGAGTTTGAT 1487

RESULT 7

US-09-866-379-6
; Sequence 6, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; CURRENT FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

Query Match 97.04; Score 1283; DB 10; Length 1901;
Best Local Similarity 99.28; Pred. No. 0;
Matches 1290; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy 1 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCGCGCATTCGCA 60
Db 188 ATGAAGCGATCTTAATCCCATTTTATCTCTCTGATTCGGTTAACCCGCGCATTCGCA 247
Qy 61 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGATTTGATTCAGTCTGATGCT 120
Db 248 TTGCGTCAGAGTGAAGCGGAGCTGAAGCTGAAGTGTGATTTGATTCAGTCTGATGCT 307
Qy 121 GTGCGTGTCTCAACCAAGGCGCAGCAACTGATGAGATGTCAACCCGAGAGCATGGCCA 180
Db 308 GTGCGTGTCTCAACCAAGGCGCAGCAACTGATGAGATGTCAACCCGAGAGCATGGCCA 367

Qy 181 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGGNGGTGGTGAAGTAAATGCCATATCTC 240
Db 368 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGGNGGTGGTGAAGTAAATGCCATATCTC 427
Qy 241 GGACATTACCAACGCCACGCTCTGTGTACCGACGAGATTGCTGGCGAAAAGGCTGCCCG 300
Db 428 GGACATTACCGCGCGACGCTCTGTGTACCGACGAGATTGCTGGCGAAAAGGCTGCCCG 487
Qy 301 CAGTGTGTGATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 360
Db 488 CAGTGTGTGATGATGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 547
Qy 361 GCTTTCGCGCGCGGCTGGGCACTGATGATGATTAACCGGATTAACCGGAGAGATACG 420
Db 548 GCTTTCGCGCGCGGCTGGGCACTGATGATGATTAACCGGATTAACCGGAGAGATACG 607
Qy 421 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGCCAACTGATTAACGCG 480
Db 608 TCCAGTCCCGATCCGTTATTTATCTCTTAATAAATCGGCGTTGCCAACTGATTAACGCG 667
Qy 481 AACGTGACTGACGAGATCTCAGACGAGGCGAGAGGCTCAATGCTGACTTACCGGCGAT 540
Db 668 AACGTGACTGACGAGATCTCAGACGAGGCGAGAGGCTCAATGCTGACTTACCGGCGAT 727
Qy 541 CGGCAAAAGCGGCTTCCGCAACTGCAAGCGGCTGCTTAATTTCCGCAATCAACTGTGCG 600
Db 728 CGGCAAAAGCGGCTTCCGCAACTGCAAGCGGCTGCTTAATTTCCGCAATCAACTGTGCG 787
Qy 601 CTTAAACGTGAGAAACAGAGCAAGAACTGTTCATTAAACGAGCATTAACATCGGAATCTC 660
Db 788 CTTAAACGTGAGAAACAGAGCAAGAACTGTTCATTAAACGAGCATTAACATCGGAATCTC 847
Qy 661 AAGGTGAGCGCGGCAATGCTCTAATTAACCGGCTGCTTAACCGGCTGCTTAACCGGCTG 720
Db 848 AAGGTGAGCGCGGCAATGCTCTAATTAACCGGCTGCTTAACCGGCTGCTTAACCGGCTG 907
Qy 721 GAGATATTTCTCCGCAACAGCAGAGGAAATGCGGAGCGGCGGAGGAGGATGATCACC 780
Db 908 GAGATATTTCTCCGCAACAGCAGAGGAAATGCGGAGCGGCGGAGGAGGATGATCACC 967
Qy 781 GATTACACAGCAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 840
Db 968 GATTACACAGCAGTGAACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTTGCTACAA 1027
Qy 841 CGCAGCGCAAGAGTTGCGCGGAGCGGCGGCAACCGGCTTATGATTTGATGATGAGCAGCG 900
Db 1028 CGCAGCGCAAGAGTTGCGCGGAGCGGCGGCAACCGGCTTATGATTTGATGATGAGCAGCG 1087
Qy 901 TTGACGCGCCATCCACGCAAAAACAGCGGATGATGATGATGATGATGATGATGATGATGAT 960
Db 1088 TTGACGCGCCATCCACGCAAAAACAGCGGATGATGATGATGATGATGATGATGATGATGAT 1147
Qy 961 TTTATTCGCGACACGATTAATCTGCAAAATCTGCGCGGCGGCACTGAGAGCTCACTGCG 1020
Db 1148 TTTATTCGCGACACGATTAATCTGCAAAATCTGCGCGGCGGCACTGAGAGCTCACTGCG 1207
Qy 1021 ACCTTCGCGGTGACGCGGATTAACACGCGCGGCGGAGTGTGATGATGATGATGATGATGAT 1080
Db 1208 ACCTTCGCGGTGACGCGGATTAACACGCGCGGCGGAGTGTGATGATGATGATGATGATGAT 1267
Qy 1081 CGTCGGCTAAGCATTAACACGAGTGTGATTCAGGTTTCGCTGCTTCACAGACTTTCAG 1140
Db 1268 CGTCGGCTAAGCATTAACACGAGTGTGATTCAGGTTTCGCTGCTTCACAGACTTTCAG 1337
Qy 1141 CAGATGCGGTATTAACACGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1200
Db 1328 CAGATGCGGTATTAACACGCCGCTGTCTATTAATACGCCGCCGAGAGAGTGAACCTGACC 1387
Qy 1201 CTGGCAGAGTGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCA 1260
Db 1388 CTGGCAGAGTGTGAAGGCGCAATGCGCAGGCGATGTTCTGTTGGCAGGTTTACGCA 1447
Qy 1261 ATCGTGAATGAGCAGCATACCGCGCTGAGTTTGAT 1300

Db 1448 ATCGTGAATGAAGCAGCATACCGCGGTGCTGATTTGTAAT 1487

RESULT 8

US-10-266-041-9
 : Sequence 9, Application US/10266041
 : Publication No.: US20030072844A1
 : GENERAL INFORMATION:
 : APPLICANT: LeI, Xingen
 : TITLE OF INVENTION: ENZYMES WITH IMPROVED PHYTASE ACTIVITY
 : FILE REFERENCE: 19603/2791
 : CURRENT APPLICATION NUMBER: US/10/266,041
 : CURRENT FILING DATE: 2002-10-07
 : PRIOR APPLICATION NUMBER: US/09/540,149
 : PRIOR FILING DATE: 2000-03-31
 : PRIOR APPLICATION NUMBER: 60/127,032
 : PRIOR FILING DATE: 1999-03-31
 : NUMBER OF SEQ ID NOS: 9
 : SOFTWARE: PatentIn Ver. 2.1
 : SEQ ID NO: 9
 : LENGTH: 1489
 : TYPE: DNA
 : ORGANISM: Escherichia coli
 US-10-266-041-9

Query Match 95.5%; Score 1263.8; DB 9; Length 1489;
 Best Local Similarity 97.9%; Pred. No. 0;
 Matches 1280; Conservative 0; Mismatches 28; Indels 0; Gaps 0;

QY 1 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAACCCCGCAATCTGCA 60
 Db 182 ATGAAAGCGATCTTAATCCCATTTTATCTCTCTGATTTCCGTTAACCCCGCAATCTGCA 241
 QY 61 TTGCGTCAGAGTGAAGCGAGTGAAGCTGAAAGTGTGATTTGTCAGTCGTCATGT 120
 Db 242 TTGCGTCAGAGTGAAGCGAGTGAAGCTGAAAGTGTGATTTGTCAGTCGTCATGT 301
 QY 121 GTGCGTGTCTCAACCAAGGCCAGCAACTGATGAGATGTCACCCAGACGATGCGCA 180
 Db 302 GTGCGTGTCTCAACCAAGGCCAGCAACTGATGAGATGTCACCCAGACGATGCGCA 361
 QY 181 ACCTGGCGGTAAACGAGGTGTGGCTGACACCGGNGGTGTGAGTAAATCGCTATCTC 240
 Db 362 ACCTGGCGGTAAACGAGGTGTGGCTGACACCGGNGGTGTGAGTAAATCGCTATCTC 421
 QY 241 GGACATTACCAAGCCAGCGTGTGAGCCGAGGATTTGCGGCAAAAAGGCGTCCCG 300
 Db 422 GGACATTACCAAGCCAGCGTGTGAGCCGAGGATTTGCGGCAAAAAGGCGTCCCG 481
 QY 301 CAGTCTGTGTCAGTCCGATTTATTTGTCGATGTGAGAGCGATACCCGTAACAGCGAA 360
 Db 482 CAGTCTGTGTCAGTCCGATTTATTTGTCGATGTGAGAGCGATACCCGTAACAGCGAA 541
 QY 361 GCGTTCGCGCGCGGCGTGGAGCACTGATGTGCAATACCGTATACCGAGCGATACG 420
 Db 542 GCGTTCGCGCGCGGCGTGGAGCACTGATGTGCAATACCGTATACCGAGCGATACG 601
 QY 421 TCCAGTCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCAACTGATACGCG 480
 Db 602 TCCAGTCCGATCCGTTATTTATCTCTTAAAACTGGCGTTTCCCAACTGATACGCG 661
 QY 481 AACGTGACGACGCGATCTCAGCAGGAGGAGGAGGATTCGTCGATTTACCGGCGAT 540
 Db 662 AACGTGACGACGCGATCTCAGCAGGAGGAGGAGGATTCGTCGATTTACCGGCGAT 721
 QY 541 CGGCAAAAGCGCGTTTCCGCACTGGAACGGGTGCTTAATTTTCCGCAATCAATTTGTC 600
 Db 722 CGGCAAAAGCGCGTTTCCGCACTGGAACGGGTGCTTAATTTTCCCAATCAATTTGTC 781
 QY 601 CTTAAACGTGAAGAAAGAGAGAAAGCTGTCAATTAAGCAGGATTAACCATGGAATCTC 660
 Db 782 CTTAAACGTGAAGAAAGAGAGAAAGCTGTCAATTAAGCAGGATTAACCATGGAATCTC 841

QY 661 AAGTGAAGCGCGACCAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCATGCTGAGC 720
 Db 842 AAGTGAAGCGCGACCAATGTCTCATTAACCGGTGCGGTAAGCTCGCATCATGCTGAGC 901
 QY 721 GAGTATTTCTCTCGCAACAGCAGAGGAAATGCGGAGCGGCGGTGGGAAAGATCTACC 780
 Db 902 GAGTATTTCTCTCGCAACAGCAGAGGAAATGCGGAGCGGCGGTGGGAAAGATCTACC 961
 QY 781 GATTACACCAAGTGAACACCTTGCTAAGTTTGCATTAACGCGCAATTTATTTGTACAA 840
 Db 962 GATTACACCAAGTGAACACCTTGCTAAGTTTGCATTAACGCGCAATTTATTTGTACAA 1021
 QY 841 CGCAGCGCAGAGTGTGCGCGACCGCGCAACCCCTTATTTGATTTGATGATGCGAGC 900
 Db 1022 CGCAGCGCAGAGTGTGCGCGACCGCGCAACCCCTTATTTGATTTGATGATGCGAGC 1081
 QY 901 TTGACGCGCCATCAGCAGCAAAAGAGCGGTATGTGATTAACCACTTCAGTACTG 960
 Db 1082 TTGACGCGCCATCAGCAGCAAAAGAGCGGTATGTGATTAACCACTTCAGTACTG 1141
 QY 961 TTTATTTGCGGACAGATTAATCTGCAAAATCTGCGCGCGCACTGAGCTCAACTG 1020
 Db 1142 TTTATTTGCGGACAGATTAATCTGCAAAATCTGCGCGCGCACTGAGCTCAACTG 1201
 QY 1021 ACCTTCCCGGTACGCGGATTAACAGCCCGCGAGTGTGATGATGTTGAACGCTG 1080
 Db 1202 ACCTTCCCGGTACGCGGATTAACAGCCCGCGAGTGTGATGATGTTGAACGCTG 1261
 QY 1081 CGTGGCTTAAGCATTAACAGCCAGTGTGATTCAGTGTGCTTCGACACTTTACAG 1140
 Db 1262 CGTGGCTTAAGCATTAACAGCCAGTGTGATTCAGTGTGCTTCGACACTTTACAG 1321
 QY 1141 CAGATGCGTATTAACAGCCGCTGTCTTAAATACGCGCGCGAGAGGTGAACACTGAC 1200
 Db 1322 CAGATGCGTATTAACAGCCGCTGTCTTAAATACGCGCGCGAGAGGTGAACACTGAC 1381
 QY 1201 CTGGCAGAGTGTGAAGAGCAAAATGCGCAGGAGCATGTTGTTGCGAGGTTTACGCA 1260
 Db 1382 CTGGCAGAGTGTGAAGAGCAAAATGCGCAGGAGCATGTTGTTGCGAGGTTTACGCA 1441
 QY 1261 ATCGTGAATGAAGCAGCATACCGCGGTGACGTTTGATCTCATAC 1308
 Db 1442 ATCGTGAATGAAGCAGCATACCGCGGTGACGTTTGATCTCATAC 1489

RESULT 9

US-10-021-723A-3
 : Sequence 3, Application US/10021723A
 : Publication No.: US20030101476A1
 : GENERAL INFORMATION:
 : APPLICANT: Short, Jay
 : APPLICANT: Mathur, Eric
 : APPLICANT: Richardson, Toby
 : APPLICANT: Robertson, Dan
 : APPLICANT: Barton, Nelson
 : TITLE OF INVENTION: Recombinant phytases and uses thereof
 : FILE REFERENCE: 112766.140 (Div-016CIP)
 : CURRENT APPLICATION NUMBER: US/10/021,723A
 : CURRENT FILING DATE: 2002-10-23
 : PRIOR APPLICATION NUMBER: US 60/255,090
 : PRIOR FILING DATE: 2000-12-12
 : NUMBER OF SEQ ID NOS: 74
 : SOFTWARE: FastSeq for Windows Version 4.0
 : SEQ ID NO: 3
 : LENGTH: 1325
 : TYPE: DNA
 : ORGANISM: Yersinia pestis
 US-10-021-723A-3

Query Match 16.5%; Score 218.8; DB 9; Length 1325;
 Best Local Similarity 52.8%; Pred. No. 9e-67;
 Matches 567; Conservative 0; Mismatches 493; Indels 13; Gaps 4;


```

; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0.
; SEQ ID NO 9
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 954-2 phytase sequence
US-10-021-723A-9

```

Query Match 12.1%; Score 159.6; DB 9; Length 1431;

Best Local Similarity 49.3%; Pred. No. 1,1e-45;

Matches 575; Conservative 0; Mismatches 565; Indels 27; Gaps 5;

```

OY 75 GCGGAGCTGAAGCTGGAAGTGTGATGTCAGTCATGTCGTGCTCCAC 134
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 168 GCAAGACTGCTGCTGACGCGTGTGTCATGTCAGCGCGGCGTCTCCGAC 227
OY 135 CAAGGCCACCACTGATGAGATGTCACCCACAGCGCATGGCCAACTGGCCGTTAA 194
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 228 GCGTCCGAGCCCGCTGAGAGCCCTGACCGCCGATCCGTCGCGCCGCGCCGTCGCC 287
OY 195 ACTGGTGGCTGACACCGCGGAGTGTGAGCTAATGCTATCTGCGCTCCAC 254
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 288 GATCGCCCACTGACCGATGTCGCGCGGCGCTGTCGACAGATGGGCGGCTACTACG 347
OY 255 CCAAGCTGTGTGATGCGAGCGATGTCGCGGAAAGAGGCTGCCGAGTCTGTCAGGT 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 348 TGAATGGCTTCGTCGCGGCGGTGTCGCGCGGCGGCGGCTCGACCGGAGAGCT 407
OY 315 CCGGATATGTCGATGTCGACGACGCTACCCGTAACAGCGGCAAGCTTCGCGCGCG 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 408 TTACGATGGGACAGCTTACCAACGCGGCGCTGTCAGCGGCGGCGGCTGCTCTCG 467
OY 375 GCTGGACCTGATGTCGATTAACCGTACATACCGACGAGATACGTCCCGATCC 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 468 CATGGCGCCGAGCGTGTGTCACACAGCATCATCGGCGCGCTCCACAGAGATCC 527
OY 435 GTTATTTAATCTCTTAAACCTGGCGTTTCCCACTGGATTAACCGAAGCTGACTGAC 494
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 528 GATCTTCACGCGATGGAATCGGGTGCATGCCAGTCGACCCCTACAGGGAAGCGCGA 587
OY 495 GATCTTCAGCAGGCGAGAGGATCATTTGCTGACTTACCGGCGCATGGCAACGGCGTT 554
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 588 CATCGAAGCCCATGCGCGGAG--GCGGCGGTGGCGACACTGGGAAGCGGTAGCGAC 645
OY 555 TCGGAACTGGAAGCGGCTGCTTAATTTCCGCAATCAAACTGTCCTTAAACGTGAGAA 614
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 646 AGCGTGACCAAGATGAGCGAGGTCTGACTACGCCCATAGCGCGGATTTGGCGAAGCAT 705
OY 615 ACAGGACGAAAGCTGTTCAATTAACGAGCATTAACATCGGAACTCAAGGTAGCGCGA 674
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 706 GCGGCGCAATGCGATGTGGCGGCCAACCAGTGTGCGAGATCAGACAGATGGCCTT 765
OY 675 CAATGTCTCATTAACCGGTGCGGTAGCTTCGATCAATGCTGACGAGATATTTCCT 734
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 766 C-ATGCCGCTTGAAGGCGCCGATGAGGAGTCTTCACCGTCTCCAGGCTTCCT 824
OY 735 GCAAGACGACAGGAGATGCC-----GAGCGCGGGTGGGGAAGATCAACAGATTACA 788
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 825 CGAAGCATGGGCGAGGCTGCGCACAGAGCAGGTTCATGGGCGCTTTCACAGTTCGCGCA 884
OY 789 CAGGTGAGACCTTGTCTAAGTTTGCATTAACGCGCAATTTTATTTGCTACAAACGACGCC 848
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 885 GGAAGTGAAGCTGTGATGAGGCGCATTAACGCCAGTTCGATGATGGCGAAGACGCC 944

```

```

OY 849 AGAGTTGCCCGAGCGCGCACCCGTTATGATTTGATCATGACGCGT----- 901
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 945 TTACATGCTACTTCGAAGGGGACGCCGATGCTGCTTCGCTGATGCGCTTGACG 1004
OY 902 -----TGACGCCCATCCACCCGAAAGAGCGTATGATGATTAACCATTCAG 955
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1005 GCGGCGTGGGCGCCACCTCCGAGCTTGCGGCTGAAGGCGCCGAGCTGCCAAGGCA 1064
OY 956 TACTGTTTA-----TTCCGAGACAGTACTTAATCTGGCAATCTGCGCGGCACTGGA 1010
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1065 CCGTCTATGTCGTGACCGCGATGACAGAACTTTCACACTGCGCGCTGCTGCA 1124
OY 1011 GCTCACTGACAGCGTTCGCGGTGCGGTAACACGCGCGGAGTGTGATGATGCTT 1070
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1125 CTGAGCTGACCGCTGCGCCGACAGCAGACAGCAGCAGCGCGCGCTGATGCTGTT 1184
OY 1071 TGAACGCTGCGCTGCTAAGCGATTAACAGCCAGTGTGATGCTGCTTCCA 1130
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1185 CTCTTGTGGCGGAGCGCTGGACAGGACGATTTGCTGCGTGAGATGGTCTATCA 1244
OY 1131 GACTTTACAGCAGATGCTGATTAACCGCGCTGCTATTAATACGCGCGCGAGAGGT 1190
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1245 GTCGATGATCAGCTTTCGCGAGCTCACCGCGCTCTCCCTGCGCGAGCGCGCTATCGCCT 1304
OY 1191 GAAACGACCCCTGGCAGGATGTGA 1217
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1305 GATCTTCCGTTGCGCGCTGTGCGCA 1331

```

RESULT 13

US-10-021-723A-7

Sequence 7, Application US/10021723A

Publication No. US20030101476A1

GENERAL INFORMATION:

APPLICANT: Short, Jay

APPLICANT: Mathur, Eric

APPLICANT: Richardson, Toby

APPLICANT: Robertson, Dan

APPLICANT: Barton, Nelson

TITLE OF INVENTION: Recombinant Phytases and Uses Thereof

FILE REFERENCE: 112766.140 (DIV-016CIP)

CURRENT APPLICATION NUMBER: US/10/021,723A

PRIOR APPLICATION NUMBER: US 60/255,090

PRIOR FILING DATE: 2000-12-12

NUMBER OF SEQ ID NOS: 74

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 7

LENGTH: 1230

ORGANISM: Rhizobium

US-10-021-723A-7

Query Match 5.5%; Score 73; DB 9; Length 1230;

Best Local Similarity 54.1%; Pred. No. 7,1e-15;

Matches 170; Conservative 0; Mismatches 141; Indels 3; Gaps 1;

```

OY 81 GCTGAAGCTGGAAGTGTGATGTCATGTCATGTCGTCGCTCCACCAAGGC 140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 93 GCTAAGCTGGAAGTGTGATGTCATGTCATGTCGTCGCTCCACCAAGGC 152
OY 141 CAGCACTGATGAGATGTCACCCAGACGATGGCCAACTGGCGCGGTAACATGCG 200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 153 GCGGCTGTGCGCGCGCTTATTTGCGGAAACATGCGCGGCTGCGATTCGG 212
OY 201 TTGGCTGACACGCGGCGTGTGAGCTAATGCGCTATCTGAGCATTAACCAAGCCAGCG 260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 213 CTCTGTGACGCGGACGCGCGCGGCGGCTTAAGCTTCTGCGCAAGCAGCCCTCTA 272
OY 261 TCTGTGAGCGGAGATGCTGCGGAAAGAGGCGTCCGCGAGTCTGCTAGTCCGAT 320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 273 TTTCGGGCGTGCAGGCGCTGTTTC---GAGGAGATGCCGCGCGGCGGACGATGTCCT 329

```

QY 321 TATTGTCAGTGCAGCAGCCTTAAACAGCGCAAGCTTCCCGCGGCTGC 380
 Db 330 CAAGCGAGCTACAAAGCGCAGATGCACCGCAGAGCTGGCGCGGCTTCAT 389
 QY 381 ACCTGACTGCGCA 394
 Db 390 GCCCGCTGCGCGA 403

RESULT 14

US-10-021-723A-5

; Sequence 5, Application US/10021723A
 ; Publication No. US20030101476A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Mathur, Jay
 ; APPLICANT: Short, Eric
 ; APPLICANT: Richardson, Toby
 ; APPLICANT: Barton, Nelson
 ; APPLICANT: Robertson, Dan
 ; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
 ; FILE REFERENCE: 112766.140 (DIV-016CIP)
 ; CURRENT APPLICATION NUMBER: US/10/021,723A
 ; PRIORITY FILING DATE: 2002-10-23
 ; PRIOR APPLICATION NUMBER: US 60/255,090
 ; NUMBER OF SEQ ID NOS: 74
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 5
 ; LENGTH: 1266
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: 953-6 phytase sequence
 ; US-10-021-723A-5.

Query Match

4.9%; Score 64.8; DB 9; Length 1266;

Best Local Similarity 52.8%; Pred. No. 6e-12; Matches 163; Conservative 0; Mismatches 143; Indels 3; Gaps 1;

QY 81 GCTGAACCTGGAAGTGTGATGTCAGTCATGTCAGTGTGCTCCCAACCAAGC 140
 Db 129 GCTGAACCTGGAAGTGTGATGTCAGTCATGTCAGTGTGCTCCCAACCAAGC 188
 QY 141 CACGCAATGATGATGATGTCACCCAGACGATGCGCAACTGCGCGTAAACTGG 200
 Db 189 GCGGCTGTCGCGCGCGTATTCGCGCAACATGCGCGCGTATTCGCG 248
 QY 201 TTGGCTGACACCGGCGGCTGTCATGTCATGTCATGTCATGTCATGTCATG 260
 Db 249 CTGCTGACCG 308
 QY 261 TCTGTCAGCGCAGATGTCGCGCAAAAGGCTGCGCGCGCGCGCGCGCGCGAT 320
 Db 309 TTTCGGGCTGCG 365
 QY 321 TATTGTCAGTGCAGCAGCGTAAACAGCGCAAGCTTCCCGCGCGCGCTGC 380
 Db 366 CAAGCGAGCTACAAAGCGCAGATGCACCGCGCGCGCGCGCGCGCGCGCTCAT 425
 QY 381 ACCTGACTG 389
 Db 426 GCCCGCTG 434

RESULT 15

US-10-184-644-312

; Sequence 312, Application US/10184644
 ; Publication No. US20030044930A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Chen, Jian
 ; APPLICANT: Desnoyers, Luc

; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Matanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3430R1C227
 ; CURRENT APPLICATION NUMBER: US/10/184,644
 ; PRIORITY FILING DATE: 2002-06-28
 ; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
 ; NUMBER OF SEQ ID NOS: 612
 ; SEQ ID NO 312
 ; LENGTH: 802
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-10-184-644-312

Query Match

2.8%; Score 36.4; DB 9; Length 802;

Best Local Similarity 5.2%; Pred. No. 0.06; Matches 32; Conservative 190; Mismatches 395; Indels 0; Gaps 0;

QY 682 TCATTAAACGGTGGTGAACCTTCGATCAATGCTGACGAGATATTCTCTGCAACA 741
 Db 112 TTPKVOFFSESGSDTVPPIVYPLRGVDSYPPQKSFMLKMHYLDKEMFMRAD 171
 QY 742 GCACAGGAGATGCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 801
 Db 172 DDVTKGDRLENFIRLSNLSSEPLFLGOTGLTTEEMKLLDPEBNFMGPGVIMREV 231
 QY 802 TTGCTAAGTTTGATACGCGCAATTTTATTTGCTACACGACCGCGAGTTCGCCG 861
 Db 232 LRNVPHIKGRLRMVTHHEVEGRVRRPAGVQVWSYEMRQLFEYENQKKYIRD 291
 QY 862 AGCG 921
 Db 292 LHSKIHQATTLHKNKMPPOYLHLSYMLSRKISLHRTIQLHRETVLMSKSYNTEIHK 351
 QY 922 AAACAGCGATGATGTCATGATTCACCATTCATGATGATTTATTCGCGGACGATCT 981
 Db 352 EDIOLGTPSPFMRQPPREBILMEWFLTKLYLSAVDGPFRGMDSAGREALDLYMQ 411
 QY 982 AATCTGCAATCTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAT 1041
 Db 412 VMEMINMANRTGRILIDFKELGYRVMYGALEYILDLLLYKKHKKMTPVVRHA 471
 QY 1042 AACAGCG 1101
 Db 472 YLAQTFPSKIOFVEHEBELDAQELAKRINOESGSLSFSLNSLTKLVPFQLPSKSEHKRPKD 531
 QY 1102 CAGTGGATTCAGTTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1161
 Db 532 KKNILIPLSGRFMEVRFMGNFEKTLIPRONKVLVLLFNSDSNDKAKOVELMDYR 591
 QY 1162 CTGTCAATTAATACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1221
 Db 592 IKYPKADMLLPVSGERSRLALEVSSQNNESLLEFCVDLVFTTEFLORCRANTVLG 651
 QY 1222 AATGCGAGGATGTTGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1281
 Db 652 QOIYFPIIFSOYDPKIVYSGKVPDSNHFATOKTGFWRNNGFGITCIYKGLDVRVGGFDV 711
 QY 1282 CCGCGCTGCGATTTGAG 1298
 Db 712 STGGWGLEVDVLEKNV 728

Search completed: June 12, 2003, 10:27:21

Job time : 244 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 08:19:44 ; Search time 3560 Seconds

(without alignments)
10815.453 Million cell updates/sec

Title: US-09-777-566a-1

Perfect score: 1323

Sequence: 1 atgaagcagcatctatcc.....atccatcaccatcactaa 1323

Scoring table:

IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_da:*
- 2: gb_htg:*
- 3: gb_in:*
- 4: gb_ov:*
- 5: gb_ov:*
- 6: gb_pat:*
- 7: gb_ph:*
- 8: gb_pl:*
- 9: gb_pr:*
- 10: gb_ro:*
- 11: gb_sts:*
- 12: gb_sy:*
- 13: gb_un:*
- 14: gb_vl:*
- 15: em_ba:*
- 16: em_fun:*
- 17: em_hum:*
- 18: em_in:*
- 19: em_mu:*
- 20: em_com:*
- 21: em_or:*
- 22: em_ov:*
- 23: em_ph:*
- 24: em_pl:*
- 25: em_pi:*
- 26: em_ro:*
- 27: em_sts:*
- 28: em_un:*
- 29: em_vl:*
- 30: em_htg_hum:*
- 31: em_htg_inv:*
- 32: em_htg_other:*
- 33: em_htg_mus:*
- 34: em_htg_pln:*
- 35: em_htg_rtd:*
- 36: em_htg_mam:*
- 37: em_htg_vrt:*
- 38: em_sy:*
- 39: em_htgo_hum:*
- 40: em_htgo_mus:*
- 41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1322	99.9	1323	6	AR108133
2	1322	99.9	1323	6	AR127818
3	1322	99.9	1323	6	AR130956
4	1322	99.9	1323	6	AX052713
5	1322	99.9	1323	6	AX338966
6	1322	99.9	1323	6	AX356566
7	1287.8	97.3	1901	6	AX356572
8	1287.8	97.3	1901	6	AX356572
9	1287.8	97.3	10784	1	AE000200
10	1286.2	97.2	18882	1	D90735
11	1286.2	97.2	1901	6	ECOAACPHB
12	1286.2	97.2	1901	6	AX356570
13	1286.2	97.2	1901	6	AX356574
14	1284.6	97.1	1901	1	ECOAACPHC
15	1284.6	97.1	1901	1	ECOAACPHD
16	1284.6	97.1	1901	1	ECOAACPHF
17	1284.6	97.1	1901	1	ECOAACPHF
18	1283	97.0	1901	1	ECOAACPHA
19	1283	97.0	1901	6	AX356571
20	1282.2	96.9	3470	6	AX042376
21	1282.2	96.9	4060	6	AX042374
22	1282.2	96.9	6116	6	AX042375
23	1282.2	96.9	6708	6	AX042373
24	1282.2	96.9	17732	6	AX042378
25	1282.2	96.9	20623	6	AX042372
26	1270.2	96.0	5421	6	AX042377
27	1246.6	94.2	10029	1	AE005292
28	1246.6	94.2	32773	1	AP002554
29	713.4	53.9	212936	2	AC020970
30	319.6	24.2	114979	2	AC020798
31	305.8	23.1	700	1	ECAPPA
32	222.8	16.8	10029	1	AE013783
33	222.8	16.8	193050	1	AJ414149
34	160.2	12.1	178373	2	AC102243
35	97.8	7.4	825	1	AE427147
36	65.6	5.0	11877	1	AE011683
37	64.8	4.9	21670	1	AE008748
38	62.4	4.7	254050	1	AL627269
39	61.6	4.7	10545	1	AE005727
40	61.4	4.6	10385	1	AE011713
41	60.6	4.6	5166	1	STU75949
42	56.8	4.3	1675	1	ECOAACPA
43	56.8	4.3	11710	1	AE000202
44	56.8	4.3	18252	1	D90737
45	55.4	4.2	13498	1	AE012175

ALIGNMENTS

RESULT 1
AR108133
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES

AR108133
Sequence 1 from patent US 6110719.
AR108133
AR108133.1 GI:12823620
Unknown.
Unclassified.
1 (bases 1 to 1323)
Kretz,K.
PhyTase
Patent: US 6110719-A 1 29-AUG-2000;
Location/Qualifiers

1323 bp
DNA
linear
PAT 14-FEB-2001

source 1. 1323
/organism="unknown"
BASE COUNT 323 a 353 c 357 g 289 t 1 others
ORIGIN

Query Match 99.98; Score 1322; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATCCGTTAAACCCGCAATCTGCA 60
|||||
1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATCCGTTAAACCCGCAATCTGCA 60

61 TTCGCTCAGATGAGCCGAGCTGAGCTGAGAAAGTGTGATTTGTCAGTCTCATG 120
|||||
61 TTCGCTCAGATGAGCCGAGCTGAGCTGAGAAAGTGTGATTTGTCAGTCTCATG 120

121 GTGCGTCTCCACCAAGGCCAGCAACTGATGCAAGATGTCACCCAGACGATGGCA 180
|||||
121 GTGCGTCTCCACCAAGGCCAGCAACTGATGCAAGATGTCACCCAGACGATGGCA 180

181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGNGTGGTGAATGATGCGCTATCTC 240
|||||
181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGNGTGGTGAATGATGCGCTATCTC 240

241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGAGATTGCTGGCAAAAGGCTGCCG 300
|||||
241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGAGATTGCTGGCAAAAGGCTGCCG 300

301 CAGCTGTGTGACGCTGGGATTTATGCTGATGTGACACGAGCTACCCGTTAAACAGCGCA 360
|||||
301 CAGCTGTGTGACGCTGGGATTTATGCTGATGTGACACGAGCTACCCGTTAAACAGCGCA 360

361 GCTTTCGCGCGCGGGCTGGACACTGACTGTGCAATTAACCCGATACCCAGGAGATGCG 420
|||||
361 GCTTTCGCGCGCGGGCTGGACACTGACTGTGCAATTAACCCGATACCCAGGAGATGCG 420

421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGGCGTTGCCAATGGATTAACGCG 480
|||||
421 TCCAGTCCCGATCCGTTATTTAATCTCTAATAAACTGGCGTTGCCAATGGATTAACGCG 480

481 AACGTGATGACGCGGATCTTCAGACAGGCGAGGAGGTCATGCTGACTTACCGGGCAT 540
|||||
481 AACGTGATGACGCGGATCTTCAGACAGGCGAGGAGGTCATGCTGACTTACCGGGCAT 540

541 CGGCAAAAGCGGCTTCCGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
|||||
541 CGGCAAAAGCGGCTTCCGAACTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600

601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGATTAACATCGGAATCTC 660
|||||
601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGATTAACATCGGAATCTC 660

661 AAGGTGAGCGCGCATGCTCTCATTAACGGGTGGGTAAGCCCTGCAATCAATCTGACG 720
|||||
661 AAGGTGAGCGCGCATGCTCTCATTAACGGGTGGGTAAGCCCTGCAATCAATCTGACG 720

721 GAGATATTTCTCTGCAACAAGCAAGGAGATGCCGAGCGGGGTGGGGAAGATCAAC 780
|||||
721 GAGATATTTCTCTGCAACAAGCAAGGAGATGCCGAGCGGGGTGGGGAAGATCAAC 780

781 GATTACACCAAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTATTTCTCAAA 840
|||||
781 GATTACACCAAGTGAACACCTTGTCTAAGTTTGCATTAACGCGCAATTTATTTCTCAAA 840

841 CGACGCGCAGAGGTGGCGGACGCGGCGCACTCCGTTATGATTTGATGATGCGACGCG 900
|||||
841 CGACGCGCAGAGGTGGCGGACGCGGCGCACTCCGTTATGATTTGATGATGCGACGCG 900

901 TTGACGCCCATCCACCGCAAAAGAGCGATGATGGTGAATTAACCACTTCACTGATG 960
|||||
901 TTGACGCCCATCCACCGCAAAAGAGCGATGATGGTGAATTAACCACTTCACTGATG 960

961 TTATTTCCGAGACAGACTACTAATCTGGCAAAATCTGGCGGCGCACTGGAGCTCAATCG 1020
|||||
961 TTATTTCCGAGACAGACTACTAATCTGGCAAAATCTGGCGGCGCACTGGAGCTCAATCG 1020

1021 ACCGTTCCCGGTACCGCGGATTAACAGCGCGGCGAGTGTGTAATGTTTGAACGCTGG 1080
|||||
1021 ACCGTTCCCGGTACCGCGGATTAACAGCGCGGCGAGTGTGTAATGTTTGAACGCTGG 1080

1081 CGTCGCGTAAACGATTAACAGCGCAGTGTGATGTTTGTGCTTCCAGACTTTACAG 1140
|||||
1081 CGTCGCGTAAACGATTAACAGCGCAGTGTGATGTTTGTGCTTCCAGACTTTACAG 1140

1141 CAGATGCGTGTAAACCGCGCTGTATTAATACCGCGCGGAGAGAGTGAACCTGAC 1200
|||||
1141 CAGATGCGTGTAAACCGCGCTGTATTAATACCGCGCGGAGAGAGTGAACCTGAC 1200

1201 CTGCGAGATGTGAAGACGAAATGCGGAGGCGATGTTGTTGGCAGGTTTACGCA 1260
|||||
1201 CTGCGAGATGTGAAGACGAAATGCGGAGGCGATGTTGTTGGCAGGTTTACGCA 1260

1261 ATCGGATGAAGCAGCGATACCGGCGTGCAGTTTGATGATCTCATACCATCACTAC 1320
|||||
1261 ATCGGATGAAGCAGCGATACCGGCGTGCAGTTTGATGATCTCATACCATCACTAC 1320

1321 TAA 1323
1321 TAA 1323

RESULT 2
AR127818 1323 bp DNA linear PAT 16-MAY-2001
LOCUS AR127818
DEFINITION Sequence 1 from patent US 6183740.
ACCESSION AR127818
VERSION AR127818.1 GI:14115480
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1323)
AUTHORS Short, J.M. and Kretz, K.A.
TITLE Recombinant bacterial phytases and uses thereof
JOURNAL Patent: US 6183740-A 1 06-FEB-2001;
FEATURES
source 1. 1323
/organism="unknown"

BASE COUNT 323 a 353 c 357 g 289 t 1 others
ORIGIN

Query Match 99.98; Score 1322; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATCCGTTAAACCCGCAATCTGCA 60
|||||
1 ATGAAGGAGATCTTAATCCATTTTATCTCTGATCCGTTAAACCCGCAATCTGCA 60

61 TTCGCTCAGATGAGCCGAGCTGAGCTGAGAAAGTGTGATTTGTCAGTCTCATG 120
|||||
61 TTCGCTCAGATGAGCCGAGCTGAGCTGAGAAAGTGTGATTTGTCAGTCTCATG 120

121 GTGCGTCTCCACCAAGGCCAGCAACTGATGCAAGATGTCACCCAGACGATGGCA 180
|||||
121 GTGCGTCTCCACCAAGGCCAGCAACTGATGCAAGATGTCACCCAGACGATGGCA 180

181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGNGTGGTGAATGATGCGCTATCTC 240
|||||
181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGGNGTGGTGAATGATGCGCTATCTC 240

241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGAGATTGCTGGCAAAAGGCTGCCG 300
|||||
241 GGACATTACCAACGCCAGCGCTGTGTAGCCGACGAGATTGCTGGCAAAAGGCTGCCG 300

QY	301	CAGCTGGTACAGGTCGCGATTATTGTCGTATGTCCAGCAGCGCTACCCGTTAAACAGGCGAA	360
Db	301	CAGCTGGTACAGGTCGCGATTATTGTCGTATGTCCAGCAGCGCTACCCGTTAAACAGGCGAA	360
QY	361	GCGTTGCGCGCGGGGTGGACACTGACTGTGCTCAATTAACCGTACATACCGAGCGAGATACG	420
Db	361	GCGTTGCGCGCGGGGTGGACACTGACTGTGCTCAATTAACCGTACATACCGAGCGAGATACG	420
QY	421	TCGAGTCCCGATCCGTTATTTAATCTCTTAATAAACCTGGCGTTTGCCAACTGATTAACGCG	480
Db	421	TCGAGTCCCGATCCGTTATTTAATCTCTTAATAAACCTGGCGTTTGCCAACTGATTAACGCG	480
QY	481	AACGTGACTGACGGGATCTCAGACAGGGCAGAGGGGTCAATTGGCTGACTTTACGGGGAT	540
Db	481	AACGTGACTGACGGGATCTCAGACAGGGCAGAGGGGTCAATTGGCTGACTTTACGGGGAT	540
QY	541	CGGCAAAACGGCGTTTCGGAACTGGAACTGGGTGCTTAATTTTCGGCAATCAACTTGTGC	600
Db	541	CGGCAAAACGGCGTTTCGGAACTGGAACTGGGTGCTTAATTTTCGGCAATCAACTTGTGC	600
QY	601	CTTAAGCTGAGAAACAGAGAAAGCGTTCATTTAACGCAAGGCAATTACCATCGGAATC	660
Db	601	CTTAAGCTGAGAAACAGAGAAAGCGTTCATTTAACGCAAGGCAATTACCATCGGAATC	660
QY	661	AAGGTGAGCGCGCAATGTCATTAACCGGTGCGTAAACCTCGCATCATGCTGCAGC	720
Db	661	AAGGTGAGCGCGCAATGTCATTAACCGGTGCGTAAACCTCGCATCATGCTGCAGC	720
QY	721	GAGATATTTCCTCTGCAACAAACAGCAGGGAATGCCGAGCGCGGGGTGGGAAGATCAAC	780
Db	721	GAGATATTTCCTCTGCAACAAACAGCAGGGAATGCCGAGCGCGGGGTGGGAAGATCAAC	780
QY	781	GATTACACACAGTGGGAACACCTCTAAGTTGTCATAACGCGCAATTTATTGGTTACAA	840
Db	781	GATTACACACAGTGGGAACACCTCTAAGTTGTCATAACGCGCAATTTATTGGTTACAA	840
QY	841	CGCACGCCAGAGTTGGCCCGCAGCGCGCCACCCCGTTATTGGATTGATCATGGCAGCG	900
Db	841	CGCACGCCAGAGTTGGCCCGCAGCGCGCCACCCCGTTATTGGATTGATCATGGCAGCG	900
QY	901	TTGACGGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTTCCCACTTCAGTACTG	960
Db	901	TTGACGGCCCCATCCACCGCAAAAACAGGCGTATGGTGTGACATTTCCCACTTCAGTACTG	960
QY	961	TTTATTGGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGAGACTCAACTGG	1020
Db	961	TTTATTGGCGGACACGATACTAATCTGGCAATCTCGCGCGCGCACTGAGACTCAACTGG	1020
QY	1021	ACGCTTCCCGGTCAACCGCGGATTAACCGCCGAGGTGGTGAACCTGGTTGAACGCTGG	1080
Db	1021	ACGCTTCCCGGTCAACCGCGGATTAACCGCCGAGGTGGTGAACCTGGTTGAACGCTGG	1080
QY	1081	CGTGGCGCTAACGATTAACGACGACGAGTTCAGGTTGCGTGGTCTTCACACTTTACAG	1140
Db	1081	CGTGGCGCTAACGATTAACGACGACGAGTTCAGGTTGCGTGGTCTTCACACTTTACAG	1140
QY	1141	CAGATGCGGTATAAAACGCGCTGTCTATTAATTAACGCCCGCCGGAGAGGTGAACCTGCAC	1200
Db	1141	CAGATGCGGTATAAAACGCGCTGTCTATTAATTAACGCCCGCCGGAGAGGTGAACCTGCAC	1200
QY	1201	CTGGCAGGATGTGAAGAGCGAAATGCGAGGGCATGTGTGCTTGGCGAGGGTTTAACGCA	1260
Db	1201	CTGGCAGGATGTGAAGAGCGAAATGCGAGGGCATGTGTGCTTGGCGAGGGTTTAACGCA	1260
QY	1261	ATGCTGAAATGAAGACAGCATACCGGCGTGCAGTTTGAGATCTCATACCATCACCATC	1320
Db	1261	ATGCTGAAATGAAGACAGCATACCGGCGTGCAGTTTGAGATCTCATACCATCACCATC	1320
QY	1321	TAA 1323	
Db	1321	TAA 1323	

[illegible]

Db		721	GAGATATTTCCTCGCAACAGCAGCAGGATGCGGAGCCGGGGTGGGAGAGATCACC	780
Oy		781	GATTTCACACCGTGGGAACACCTTGCTAAAGTTTGATATAACGGCGAAATTTATTGCTTACAA	840
Db		781	GATTCACACCGAGTGGGAACACCTTCTAAAGTTTGATATAACGGCGAAATTTATTGCTTACAA	840
Oy		841	CGCAGCCGAGAGGTGGCCCGGAGCGCGCCACCCCGTATTGGATTGATCATGCGACGC	900
Db		841	CGCAGCCGAGAGGTGGCCCGGAGCGCGCCACCCCGTATTGATGATTTGATATGCGACGC	900
Oy		901	TTGAGCCGCCCATTCACCGCAAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG	960
Db		901	TTGAGCCGCCCATTCACCGCAAAAACAGCGGTATGGTGTGACATTACCCACTTCAGTACTG	960
Oy		961	TTTATTGCGCGACACGATACTAATCTGGCAAAATCTCGGCGCGGCACTGAGACTCAACTGG	1020
Db		961	TTTATTGCGCGACACGATACTAATCTGGCAAAATCTCGGCGCGGCACTGAGACTCAACTGG	1020
Oy		1021	ACGGTTCCCGGTCACGCGGATTAACACGCGCGCAGGTGGTGAACGTGTTTGAACGCTGG	1080
Db		1021	ACGGTTCCCGGTCACGCGGATTAACACGCGCGCAGGTGGTGAACGTGTTTGAACGCTGG	1080
Oy		1081	CGTGGCGCTAACCGATACAGCCAGCTGATAGTTTCGCTGCTTCCTCCAGACTTTACAG	1140
Db		1081	CGTGGCGCTAACCGATACAGCCAGCTGATAGTTTCGCTGCTTCCTCCAGACTTTACAG	1140
Oy		1141	CAGATGCGGTGTAATAAACGCCGCTGTCAATTAATTAACGCCGCCGGAAGAGGTGAACCTGACC	1200
Db		1141	CAGATGCGGTGTAATAAACGCCGCTGTCAATTAATTAACGCCGCCGGAAGAGGTGAACCTGACC	1200
Oy		1201	CTGGCAGAGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTGCGCAGGTTTAAACGCA	1260
Db		1201	CTGGCAGAGATGTGAAGAGCGAAATGCGCAGGGCATGTGTTGCGCAGGTTTAAACGCA	1260
Oy		1261	ATCGTGAATGAAGACAGCATACCGCGCGTGCAGTTTGAGATCTCATACCATCACCATTAC	1320
Db		1261	ATCGTGAATGAAGACAGCATACCGCGCGTGCAGTTTGAGATCTCATACCATCACCATTAC	1320
Oy		1321	TAA 1323	
Db		1321	TAA 1323	
RESULT 4				
LOCUS	AX052713	1323 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO0071728.			
ACCESSION	AX052713			
VERSION	AX052713.1	GI:12226903		
KEYWORDS				
SOURCE				
ORGANISM	Escherichia coli.			
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (bases 1 to 1323)			
AUTHORS	Short,J.M. and Kreitz,K.A.			
TITLE	Recombinant bacterial phytases and uses thereof			
JOURNAL	Patent: WO 0071728-A 1 30-NOV-2000; Diversa Corporation (US)			
FEATURES				
source	1..1323			
	/organism="Escherichia coli"			
	/db_xref="taxon:562"			
	1..1323			
	/note="unnamed protein product"			
	1/codon_start=1			
	/transl_table=1			
	/protein_id="CAC21810.1"			
	/db_xref="GI:12226904"			
	/translation="MKAILIPFLSLILPLTPQSAFASQSEPIKIESVIYISRHGVAP			
	TKATIQMODVTPDAMPTMPYKLGWITPRGGLNITLGHYQRLVADGLAKKCPQSS			
	GVAVIADVVERTRKTEAPVAGIAPCAITVHQATQSSDPLFNPLKGVGCLDNA			

Query Match	99.9%;	Score 1322;	DB 6;	Length 1323;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 1323;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
ORIGIN	323 a	353 c	357 g	289 t
				1 others
QY	1	ATGAAAGCGATCTTAATCCATTTTATCTCTTCGTGATTCGGTTAACCCCGCAATCTGCA	60	
DB	1	ATGAAAGCGATCTTAATCCATTTTATCTCTTCGTGATTCGGTTAACCCCGCAATCTGCA	60	
QY	61	TTCCGCTCAGATGACCCGCGAGCTGAACTGGAAAGTGGTGATTTGCATGCTGATCAGT	120	
DB	61	TTCCGCTCAGATGACCCGCGAGCTGAAAGTGGTGATTTGCATGCTGATGCTGATCAGT	120	
QY	121	GTGCGTGCTCCAAACCAAGGCGACGCACTGATGAGATGTCCACCCGACAGCATGGCGCA	180	
DB	121	GTGCGTGCTCCAAACCAAGGCGACGCACTGATGAGATGTCCACCCGACAGCATGGCGCA	180	
QY	181	ACCTGGCCGCTAAACCTGGTTGGCTGACACCCGCGNGTGGTGAGCTTAATGGCTATCTC	240	
DB	181	ACCTGGCCGCTAAACCTGGTTGGCTGACACCCGCGNGTGGTGAGCTTAATGGCTATCTC	240	
QY	241	GGACATTTACCAAGCGCACGCTCTGTGATGCCAGCGATTTGTGGGAAAAAAGGCTGCCCG	300	
DB	241	GGACATTTACCAAGCGCACGCTCTGTGATGCCAGCGATTTGTGGGAAAAAAGGCTGCCCG	300	
QY	301	CAGTCTGGTCAAGTCCGATTAATTCCTGATGTCAGACGAGCTACCCGTAACAGCGCA	360	
DB	301	CAGTCTGGTCAAGTCCGATTAATTCCTGATGTCAGACGAGCTACCCGTAACAGCGCA	360	
QY	361	GCGTTTCGCGCGGGGCTGGACCTGACGTGCATTAACCCGTAACCCAGCAGATACG	420	
DB	361	GCGTTTCGCGCGGGGCTGGACCTGACGTGCATTAACCCGTAACCCAGCAGATACG	420	
QY	421	TCGACGTCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGCCAATCGGATACGCG	480	
DB	421	TCGACGTCCGATCCGTTATTTAATCTCTTAAAACTGGCGTTTGCCAATCGGATACGCG	480	
QY	481	AACGTGATGACGGATCCGATCAGACAGGCGACGAGGATCAATTGCTGATTAACGGGAT	540	
DB	481	AACGTGATGACGGATCCGATCAGACAGGCGACGAGGATCAATTGCTGATTAACGGGAT	540	
QY	541	CGGCAAAACGGGCTTTCGCGAATCGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC	600	
DB	541	CGGCAAAACGGGCTTTCGCGAATCGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTGC	600	
QY	601	CTTAAGCGTGAAGAACGAGACGAAAGCTGTTCAATTAACGCAGCGATTAACCTCGGAATC	660	
DB	601	CTTAAGCGTGAAGAACGAGACGAAAGCTGTTCAATTAACGCAGCGATTAACCTCGGAATC	660	
QY	661	AAGGTGACGCGCGCAATGTCTCAATTAACGGTGGGTAACCTTCGCATCAATCTGTACG	720	
DB	661	AAGGTGACGCGCGCAATGTCTCAATTAACGGTGGGTAACCTTCGCATCAATCTGTACG	720	
QY	721	GAGATATTTCCTCTGCAACAAGCACAGGGAATGCCGACCGGGGTGGGGAAGGATCAC	780	
DB	721	GAGATATTTCCTCTGCAACAAGCACAGGGAATGCCGACCGGGGTGGGGAAGGATCAC	780	
QY	781	GATTCACACAGTGGGAACCTTGTCTAAGTTTGATTAACGGCGAATTTTATTCTCTACA	840	
DB	781	GATTCACACAGTGGGAACCTTGTCTAAGTTTGATTAACGGCGAATTTTATTCTCTACA	840	
QY	841	CGCAGCGCAGAGGTGGCCCGCAGCGCGCCACCCGTTATTTGATTTGATCATGTGGACGG	900	
DB	841	CGCAGCGCAGAGGTGGCCCGCAGCGCGCCACCCGTTATTTGATTTGATCATGTGGACGG	900	
QY	901	TTGACGCCCATTCACCGCAAAAACAGCGGTATGTGTGATCAATTAACCATTTCAGTACTG	960	

[illegible]

RESULT 5

LOCUS	AX338966	1323 bp	DNA
DEFINITION	Sequence 1 from Patent WO0189317.	linear	PAT 09-JAN-2002
ACCESSION	AX338966		
VERSION	AX338966.1	GI:18129102	
KEYWORDS			
SOURCE	Escherichia coli.		
ORGANISM	Escherichia coli		

DEPARTMENT

AUTHORS Short, J.M., Kretz, K.A. and O'Donoghue, E
TITLE Dietary aids and methods of use thereof
JOURNAL Patent: WO 0189317-A 1 29-NOV-2001;

FEATURES	Location/Qualifiers
source	1. .1323

CDS

```

/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD20553.1"
/db_xref="GI:18129103"
/translation="MKALILPLFLSLPLIPNPOSAPFQSEPELKLSEVYIARHGRVAP
TKRAQLMODVTPDAMPMPWPKLGMVLRPGGELIAYLGHYORRRLVADGLAKKCPQS
GQVAILADVDERTRKTCSEAPAAGLAPPCATTYHQADTSSDPLFENLKTGYCOLDNA
NVDAILSRAGGSTADFTGHNQIAFLRELERVANPOSNLCKRQDEQSLTQALPS
ELKASADNVSLTGAVSLASMLTEILFELQKAGQEMPEMGRITDSHOMNTLSLHNAQF
YLQRTPEVARSRATPLDLILMALTPKQAPQOKQGVTLPLPSVLYIGAHNTLANLGGK
ALELNTMLPGQPDNTPPGGELVFERMRRLDNSQNSOIVSLVFQTLQMDRKTPLSLMTY
PPGKGVKTLACGERRNAQGMCSLAGFPQIYNENAIRPCSLSRHHHHHH"
BASE COUNT      333 a      357 g      289 t      1 others
ORIGIN

```

Query Match	99.9%	Score 1322	DB 6	Length 1323
Best Local Similarity	100.0%	Pred. No. 0		
Matches 1323	Conservative 0	Mismatches 0	Indels 0	Gaps 0

OY	1	TTGAAAGGAGTCTTAATCCCATTTTATCTCTCGATTGCGGTTAAACCGGCATCTGCA	60
Db	1	ATGAAGGAGTCTTAATCCCATTTTATCTCTCGATTGCGGTTAAACCGGCATCTGCA	60
OY	61	TTGCGTCAGATGAGCCCGAGCTGAGGTGGAAAAGTGTGTGATTTGTCATGTCATGCT	120
Db	61	TTGCGTCAGATGAGCCCGAGCTGAGGTGGAAAAGTGTGTGATTTGTCATGTCATGCT	120
OY	121	GTGCGTGTCCAAACCAAGGCCACGCACTGATGCAAGATGTGACCCAGACGATGGCA	180
Db	121	GTGCGTGTCCAAACCAAGGCCACGCACTGATGCAAGATGTGACCCAGACGATGGCA	180
OY	181	ACGCGCCCGGTAACCTGGGTGGTGACACCCGNGGTGGTGAGCTAATGGCCATATCTC	240
Db	181	ACGCGCCCGGTAACCTGGGTGGTGACACCCGNGGTGGTGAGCTAATGGCCATATCTC	240
OY	241	GGACATTTACCAACGCCACGTCGTGGTACCCAGCGATTGTGGCAAAAAGGCGCCG	300
Db	241	GGACATTTACCAACGCCACGTCGTGGTACCCAGCGATTGTGGCAAAAAGGCGCCG	300
OY	301	CAGTGTGGTCAAGTCCGATTAATTTGCTGATGCGAGCGTACCCGTAAAAACAAGCGAA	360
Db	301	CAGTGTGGTCAAGTCCGATTAATTTGCTGATGCGAGCGTACCCGTAAAAACAAGCGAA	360
OY	361	GCGTTCCGCCGCCGGCGCTGGCACCTGACTGTGCAATACCGTACATACCCAGCGAGATACG	420
Db	361	GCGTTCCGCCGCCGGCGCTGGCACCTGACTGTGCAATACCGTACATACCCAGCGAGATACG	420
OY	421	TCCAGTCCCGAGTCCGTTATTTAATCCTCTAATAAACTGGCGTTTGCCAACTGGATACCGG	480
Db	421	TCCAGTCCCGAGTCCGTTATTTAATCCTCTAATAAACTGGCGTTTGCCAACTGGATACCGG	480
OY	481	AACGTGACTGACGCGCATCTCAGACGAGCGAGAGGTCATTTGCTGACTTTAACGGGCAT	540
Db	481	AACGTGACTGACGCGCATCTCAGACGAGCGAGAGGTCATTTGCTGACTTTAACGGGCAT	540
OY	541	CGGCAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAACTTTGTC	600
Db	541	CGGCAACGGCGTTTCGCGAACTGGAAACGGGTGCTTAATTTCCGCAATCAACTTTGTC	600
OY	601	CTTAAGCTGAAACAGAGGAAAGCTGTTCATTAACGACGGCATTCACATCGGAACTC	660
Db	601	CTTAAGCTGAAACAGAGGAAAGCTGTTCATTAACGACGGCATTCACATCGGAACTC	660
OY	661	AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGCGTAAACCTTCGATCAATGCTGACG	720
Db	661	AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGCGTAAACCTTCGATCAATGCTGACG	720
OY	721	GAGATATTTCTCTGCAACAAAGCACAGGGAATGCCGAGCCGGGTGGGAAGATCAC	780
Db	721	GAGATATTTCTCTGCAACAAAGCACAGGGAATGCCGAGCCGGGTGGGAAGATCAC	780
OY	781	GATTCACACCAAGTGGAAACCTGCTAGTTGCTAATACGGGCAATTTATTTGCTACAA	840
Db	781	GATTCACACCAAGTGGAAACCTGCTAGTTGCTAATACGGGCAATTTATTTGCTACAA	840
OY	841	CGCAGCGCAGAGGTTGCCCGACGCCGCCACCCGTTATTTGATTTGATGATGGCACG	900
Db	841	CGCAGCGCAGAGGTTGCCCGACGCCGCCACCCGTTATTTGATTTGATGATGGCACG	900
OY	901	TTTGACGCCCATCCACCGCAAAAAGCGCGTATGGTGTGACATTAACCATTTCACTACTG	960
Db	901	TTTGACGCCCATCCACCGCAAAAAGCGCGTATGGTGTGACATTAACCATTTCACTACTG	960
OY	961	TTTATTTGCCGGAACGATTAATCTGTGCAATCTCGCGCGCGCACCTGAGACTCAACTGG	1020
Db	961	TTTATTTGCCGGAACGATTAATCTGTGCAATCTCGCGCGCGCACCTGAGACTCAACTGG	1020
OY	1021	ACGCTTCCCGGTACCGCGGATTAACACGCCCGCAGGTGTGTAAGTGTGTTGAACGCTGG	1080
Db	1021	ACGCTTCCCGGTACCGCGGATTAACACGCCCGCAGGTGTGTAAGTGTGTTGAACGCTGG	1080
OY	1081	CGTCCGCTAAGCGATACAGCCACGACTGATTCAGTTTCGCTGCTTCACAGACTTTACAG	1140


```

Db      1081 CGTGGGGTAAACGATTAACAGCCAGTGATTCAGTTTGGTCTTCCAGACTTTACAG 1140
Oy      1141 CAGATGGCGTAAAGCCGCGTGCATTAATACGCCGCCGAGAGAGTGAATGACC 1200
Db      1141 CAGATGGCGTAAAGCCGCGTGCATTAATACGCCGCCGAGAGAGTGAATGACC 1200
Oy      1201 CTGGCAGAGATGTAAGAGCGAATGCGCAGGCGCATGTGTTGGCGAGTTTACGCA 1260
Db      1201 CTGGCAGAGATGTAAGAGCGAATGCGCAGGCGCATGTGTTGGCGAGTTTACGCA 1260
Oy      1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATAC 1320
Db      1261 ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTGAGATCTCATCACCATCACCATAC 1320
Oy      1321 TAA 1323
Db      1321 TAA 1323
Oy      1321 TAA 1323
Db      1321 TAA 1323

RESULT 6
AX356566 1323 bp DNA Linear PAT 06-FEB-2002
LOCUS     Sequence 1 from Patent WO0190333.
ACCESSION AX356566
VERSION   AX356566.1 GI:18621053
KEYWORDS  .
SOURCE    Escherichia coli.
          Escherichia coli.
          Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
          Escherichia.
REFERENCE 1
AUTHORS   Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrell, J.B. and
          O'Donoghue, E.
          Recombinant bacterial phylases and uses thereof
          Patent: WO 0190333-A 1 29-NOV-2001;
          DIVERSA CORPORATION (US)
FEATURES
   source          1..1323
                   Location/Qualifiers
                   1..1323
                   /organism="Escherichia coli"
                   /db_xref="taxon:562"
                   1..1323
                   /note="unnamed protein product"
                   /codon_start=1
                   /transl_table=1
                   /protein_id="CAD23025.1"
                   /db_xref="GI:18621054"
                   /translation="MKAILIPFLSLIPFQSAPOSEPELKIYESVIVSRHGVRAP
                   TKATQMODVTPDAMPMPYKILGWLTPRGGLIAYLGHYQKORIVADSLAKKGPQS
                   GQVATIDVDERTKRTGBAFAGLAPDCATVHTQADTSSPDLPNPLKTVCOQLDNA
                   NWTDAIISRAGSIADEFTGHRQTAFLRELEAVLPQSMCLKREKODSCSLTQALPS
                   ELKVSADNVSLTGAVSLASMLTEIFLLQOAGMEPGMGRITDSHOWTSLHNAQF
                   YLQRTPEVARSRAFLDLIMALTLPHPPOKQAGVLTPTSVLEIAGHDNINLIGG
                   ALEMTLPGQDPTPGGLVFERMRRLSDNSOMIOVSLFQTLQQRDKTPIPLSINT
                   PPGREYKTLIACGERRNAGMCKSLAGFTQIVAEARIPACSLSHHHHH"
BASE COUNT  323 a 353 c 357 g 289 t 1 others
ORIGIN
Query Match      99.9%; Score 1322; DB 6; Length 1323;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1323; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGNGTGGTGAAGCTAATGCGCTATCTC 240
Db      181 ACCTGGCCGGTAAACTGGGTGGCTGACACCGCGNGTGGTGAAGCTAATGCGCTATCTC 240
Oy      241 GGACATTTACCAACGCCACGCTGTGTAGCCGACGATTTGTCGGCAAAAAGGCTGCCG 300
Db      241 GGACATTTACCAACGCCACGCTGTGTAGCCGACGATTTGTCGGCAAAAAGGCTGCCG 300
Oy      301 CAGTGTGTCAGGTGCGGCTTATTTGCTGATGTCGAGAGCGATCCGTAACAGCGCA 360
Db      301 CAGTGTGTCAGGTGCGGCTTATTTGCTGATGTCGAGAGCGATCCGTAACAGCGCA 360
Oy      361 GCGTTGCGCGCGCGGCGTGGCAGACCTGACATGTGCAATACCGTACATACCGAGCAGATACG 420
Db      361 GCGTTGCGCGCGCGGCGTGGCAGACCTGACATGTGCAATACCGTACATACCGAGCAGATACG 420
Oy      421 TCAGATCCCGATCCGTTATTTATCTCTTAAAAACTGGCGTTTGGCAACTGATTAACCG 480
Db      421 TCAGATCCCGATCCGTTATTTATCTCTTAAAAACTGGCGTTTGGCAACTGATTAACCG 480
Oy      481 AACGTGACTGACGCGATCCTCAGAGGCGAGGAGGTCGAATTCGTGACTTTACCGGGCAT 540
Db      481 AACGTGACTGACGCGATCCTCAGAGGCGAGGAGGTCGAATTCGTGACTTTACCGGGCAT 540
Oy      541 CGGCAAAAGCGCGTTTCGGGAACGAGCGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
Db      541 CGGCAAAAGCGCGTTTCGGGAACGAGCGGTCTTAATTTTCCGCAATCAAACTTGTGC 600
Oy      601 CTTAAACGTGAGAAACAGAGAGCAAGCTGTTCATTAAAGCGAGCATTAACCGGAACTC 660
Db      601 CTTAAACGTGAGAAACAGAGAGCAAGCTGTTCATTAAAGCGAGCATTAACCGGAACTC 660
Oy      661 AAGGTGACGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCTGCAATATGCTGACG 720
Db      661 AAGGTGACGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCTGCAATATGCTGACG 720
Oy      721 GAGATATTTCTCTGCAACACAGCAGGATCCGGAACCGCGGTGGGGAAGATCAC 780
Db      721 GAGATATTTCTCTGCAACACAGCAGGATCCGGAACCGCGGTGGGGAAGATCAC 780
Oy      781 GATTCACACACAGTGGAAACACCTTGTGATTAACCGGCAATTTATTTGCTACAA 840
Db      781 GATTCACACACAGTGGAAACACCTTGTGATTAACCGGCAATTTATTTGCTACAA 840
Oy      841 CGCAGCGCAGAGGTTGCCCGCAGCGCCGACCCGCTTATTTGATTTGATCAATGCGACG 900
Db      841 CGCAGCGCAGAGGTTGCCCGCAGCGCCGACCCGCTTATTTGATTTGATCAATGCGACG 900
Oy      901 TTGACGCCCCATCCACCCGCAAAACAGGCGTATGCTGTGACATTTACCACTTCACTACTG 960
Db      901 TTGACGCCCCATCCACCCGCAAAACAGGCGTATGCTGTGACATTTACCACTTCACTACTG 960
Oy      961 TTTATATGCGGACAGATACCTAATCTGCAATTCGCGCGCGGCACTGGAGCTCAACTGG 1020
Db      961 TTTATATGCGGACAGATACCTAATCTGCAATTCGCGCGCGGCACTGGAGCTCAACTGG 1020
Oy      1021 ACAGTTCCCGGTCACCGGATTAACAGCGCGCAGGCTGATGAAGCTGTTTGAAGCTGCG 1080
Db      1021 ACAGTTCCCGGTCACCGGATTAACAGCGCGCAGGCTGATGAAGCTGTTTGAAGCTGCG 1080
Oy      1081 CGTGGCGTAAAGGATTAACAGCAGGATTCAGTTTGGTGGTGTCTTCACAGACTTTACAG 1140
Db      1081 CGTGGCGTAAAGGATTAACAGCAGGATTCAGTTTGGTGGTGTCTTCACAGACTTTACAG 1140
Oy      1141 CAGATGCGGTGATTAAGCCGCGTGCATTAATACGCCGCCGAGAGAGTGAATGACC 1200
Db      1141 CAGATGCGGTGATTAAGCCGCGTGCATTAATACGCCGCCGAGAGAGTGAATGACC 1200
Oy      1201 CTGGCAGAGATGTAAGAGCGAATGCGCAGGCGCATGTGTTGGCGAGTTTACGCA 1260
Db      1201 CTGGCAGAGATGTAAGAGCGAATGCGCAGGCGCATGTGTTGGCGAGTTTACGCA 1260

```


RESULT 7	ECOAPPAA	1901 bp	DNA	linear	BCT 04-APR-2002
LOCUS	Escherichia coli periplasmic phosphoanhydride phosphohydrolase				
DEFINITION	(appa) gene, complete cds.				
ACCESSION	M58708				
VERSION	M58708.1				
KEYWORDS	GI:145283				
SOURCE	Escherichia coli.				
ORGANISM	Escherichia coli.				
REFERENCE	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
AUTHORS	1 (bases 1 to 1901)				
TITLE	Dassa,J., Marck,C. and Boquet,P.L.				
JOURNAL	The complete nucleotide sequence of the Escherichia coli gene appa				
MEDLINE	reveals significant homology between pH 2.5 acid phosphatase and				
PUBMED	glucose-1-phosphatase				
REFERENCE	J. Bacteriol. 172 (9), 5497-5500 (1990)				
AUTHORS	J. Bacteriol. 172 (9), 5497-5500 (1990)				
TITLE	2 (sites)				
JOURNAL	Ostani,K., Harns,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van				
MEDLINE	Etten,R.L.				
PUBMED	Overexpression, site-directed mutagenesis, and mechanism of				
REFERENCE	Escherichia coli acid phosphatase				
AUTHORS	J. Biol. Chem. 267 (32), 22830-22836 (1992)				
TITLE	1429631				
JOURNAL	3 (sites)				
MEDLINE	Ostani,K. and Van Etten,R.L.				
PUBMED	Asp304 of Escherichia coli acid phosphatase is involved in leaving				
REFERENCE	group protonation				
AUTHORS	J. Biol. Chem. 268 (28), 20778-20784 (1993)				
TITLE	94012612				
JOURNAL	8407904				
MEDLINE					
PUBMED					
FEATURES					
SOURCE	Location/Qualifiers				
gene	1..1901				
RBS	/organism="Escherichia coli"				
CDS	/db_xref="taxon:562"				
	3..1522				
	/gene="appa"				
	3..8				
	/gene="appa"				
	16..108				
	/gene="appa"				
	/note="ORF"				
	/codon_start=1				
	/transl_table=1				
	/product="unknown"				
	/protein_id="AA072085.1"				
	/db_xref="GI:145284"				
	/translation="MTYLMFVGILMCSSTVYVWLDPLRS"				
	118..149				
	/gene="appa"				
	133..138				
	/gene="appa"				
	156..161				
	/gene="appa"				
	172..177				
	/gene="appa"				
	188..1486				
	/gene="appa"				
	/standard_name="acid phosphatase"				
	/codon_start=1				

		/transl_table=1	/product="periplasmic phosphoanhydride phosphohydrolase"
		/protein_id="AA72086.1"	
		/db_xref="gi:145285"	
		/translation="MKAILIPFLSLILPLTPQSAFQAQSEPEIKLESVIVRHGVRAI	
		TGAIVLMDDVPDAMPWPVKIGMLTTPGGELIAVLGHYORLADLANKGQPQQA	
		KVAITADVDERTKRTGEAFAGLAPOCAITVHQADISSPDLEFNPLKTGVCOLDNN	
		NVTNALISRAGSGSTADFTGHQGTAFRELERLVNFPQSMLCKAREQDSCTIQALP	
		ELKVSADVLSLFGAVSLASMLTEIFLLDAQGMPEPGGRITDSHOWNTLISLNAO	
		YLLOKTPENVARSRAPIDLIDLTALTPRPPOKAYGLPTPSVFIAHDYNLANLG	
		ALENMTLPQGPDPNTPPGSELVEFERMRISNSQWIOVSLVFOTLOQRDTPSLINT	
		PGEKFKTLIACCEERNAGMCSLAGTFQTVNEARIPACSL"	
		188..253	
sig_peptide		/gene="appa"	
		/gene="appa"	
		254..1483	
mat_peptide		/gene="appa"	
		/product="periplasmic phosphoanhydride phosphohydrolase"	
		299..301	
variation		/gene="appa"	
		/standard_name="H303A mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[2]	
		299..301	
variation		/gene="appa"	
		/standard_name="R16A mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[2]	
		299..301	
variation		/gene="appa"	
		/standard_name="R92A mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[2]	
		299..301	
variation		/gene="appa"	
		/standard_name="H17N mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[2]	
		1163..1165	
variation		/gene="appa"	
		/standard_name="D304A mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[3]	
		/phenotype="Asp 304 changed to Ala"	
		/replace="gct"	
		1163..1165	
variation		/gene="appa"	
		/standard_name="D304E mutant"	
		/note="created by site directed mutagenesis"	
		/citation=[3]	
		/phenotype="asp 304 changed to Glu"	
		/replace="gag"	
		1497..1522	
terminator		/gene="appa"	
		/note="rho-independent terminator"	
		BASE COUNT 474 a 499 c 499 g 428 t 1 others	
ORIGIN			
Query Match		97.3% ; Score 1287.8; DB 1; Length 1901;	
Best Local Similarity		99.5%; Pred. No. 0;	
Matches 1293; Conservative		0; Mismatches 7; Indels 0; Gaps 0;	
Oy	1 ATGAAAGCATCTTAATCCATTTCCTTGATTCGTTAACCCGCATCTGCA	60	
Db	188 ATGAAAGGATCTTAATCCATTTTATCTCTTGATTCGTTAACCCGCATCTGCA	247	
Oy	61 TTGCGTAGAGTGAACCGGAGCTGAAGCTGAAAATGTGGATGTCTAGTGTATGTT	120	
Oy	248 TTGCGTAGAGTGAACCGGAGCTGAAGCTGAAAATGTGGATGTCTAGTGTATGTT	307	

QY	121	GTGGTGTCTCCAAACCAAGCCGCAACGTAATGTCAGATGTCACCCGACGACATGGCCA	180
Db	308	GTGGTGTCTCCAAACCAAGCCGCAACGTAATGTCAGATGTCACCCGACGACATGGCCA	367
QY	181	ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGNGGTGGTGAAGCTAATGCCATATCTC	240
Db	368	ACCTGGCCGGTAAACTGGGTTGGCTGACACCGCGGNGGTGGTGAAGCTAATGCCATATCTC	427
QY	241	GGACATTTACCAAGCCGACGCTGCTGGTACCGACGGAATTGCTGGCGGAAAAAGGCGTGGCCG	300
Db	428	GGACATTTACCAAGCCGACGCTGCTGGTACCGACGGAATTGCTGGCGGAAAAAGGCGTGGCCG	487
QY	301	CAGTCCTGGTCAAGTCCGCGATTATTTGCTGATGTCGACGAGCGCTACCCGTTAAACAGCGCAA	360
Db	488	CAGTCCTGGTCAAGTCCGCGATTATTTGCTGATGTCGACGAGCGCTACCCGTTAAACAGCGCAA	547
QY	361	GCGTTTCGCGCGCGGGGTGGGACGCTGACTGTGCAATAACCGTACATACCAGGACGATACG	420
Db	548	GCGTTTCGCGCGCGGGGTGGGACGCTGACTGTGCAATAACCGTACATACCAGGACGATACG	607
QY	421	TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGCGTTCGGCACTGGATTAACCG	480
Db	608	TCCAGTCCCGATCCGTTATTTAATCTCTTAAAAACTGCGTTCGGCACTGGATTAACCG	667
QY	481	AACTGACTGACGCGATCTCTACGACGGCGAGAGGGTCAATTTGCTGACTTACCGGGCAT	540
Db	668	AACTGACTGACGCGATCTCTACGACGGCGAGAGGGTCAATTTGCTGACTTACCGGGCAT	727
QY	541	CGGCAAAAGCGGTTTCGCGAACTGGMAGGGGTCCTAATTTTCGCAATCAAACTTGTCG	600
Db	728	CGGCAAAAGCGGTTTCGCGAACTGGMAGGGGTCCTAATTTTCGCAATCAAACTTGTCG	787
QY	601	CTTAAACGTGAGAAAACAGAGCAAGACGTCGTTCAATTAACGACAGCAATTACATCGGAACCTC	660
Db	788	CTTAAACGTGAGAAAACAGAGCAAGACGTCGTTCAATTAACGACAGCAATTACATCGGAACCTC	847
QY	661	AAGGTGACGCGGCAATGTCGTAATTAACCGGTGGGTTAAACCTTCGATCAATGTTGACG	720
Db	848	AAGGTGACGCGGCAATGTCGTAATTAACCGGTGGGTTAAACCTTCGATCAATGTTGACG	907
QY	721	GAGATATTTCTCTGCAACAAGACACAGGAATGCGGACCGGGGTGGGGAAGATCAACC	780
Db	908	GAGATATTTCTCTGCAACAAGACACAGGAATGCGGACCGGGGTGGGGAAGATCAACC	967
QY	781	GATTTCACACACAGTGGGACACCTTGCTGAATTTGCAATTAACGCGCAATTTATTTGCTACAA	840
Db	968	GATTTCACACACAGTGGGACACCTTGCTGAATTTGCAATTAACGCGCAATTTATTTGCTACAA	1027
QY	841	CGCACGCCAGAGGTTGCCCGGACGCGCCACCCCGTATTTGGAATTCATGATGGCACG	900
Db	1028	CGCACGCCAGAGGTTGCCCGGACGCGCCACCCCGTATTTGGAATTCATGATGGCACG	1087
QY	901	TTGACGCCCATTCACACGCAAAAACAGGCGATGGTGTGACATTAACCCACTTCAGTACTG	960
Db	1088	TTGACGCCCATTCACACGCAAAAACAGGCGATGGTGTGACATTAACCCACTTCAGTACTG	1147
QY	961	TTTATTCGCGGACACGATCAATCAATCTGGCAAAATCTCGCGCGCGGCACTGAGACTCAACTCG	1020
Db	1148	TTTATTCGCGGACACGATCAATCAATCTGGCAAAATCTCGCGCGCGGCACTGAGACTCAACTCG	1207
QY	1021	ACGCTTCCCGGTCACGCGGATTAACAGCGCGCGAGGTGTGAACAGTGTTTGAACGCTGG	1080
Db	1208	ACGCTTCCCGGTCACGCGGATTAACAGCGCGCGAGGTGTGAACAGTGTTTGAACGCTGG	1267
QY	1081	CGTGGGCTAAGCGATTAACAGCCAGTGGATTTCAGGTTTCGCTGGCTTCACAGCTTTACAG	1140
Db	1268	CGTGGGCTAAGCGATTAACAGCCAGTGGATTTCAGGTTTCGCTGGCTTCACAGCTTTACAG	1327
QY	1141	CAGATGGGTGATTAAGCGCGCTGCTCAATTAATACGCGCGCGGAGAGGTAAACTGAC	1200
Db	1328	CAGATGGGTGATTAAGCGCGCTGCTCAATTAATACGCGCGCGGAGAGGTAAACTGAC	1387

OY		1201	CTGCAGATGTGAAGAACCAGAAATGCCGACGGCATGTCTCGTTGGCAGATTTCACCAA	1260
Dd		1398	CTGCAGATGTGAAGAAGCACAAAATCCGACGGCATGTCTCGTTGGCAGATTTCACCAA	1447
OY		1261	ATCGTAATGAAGCAACGCATACCGGCGTGCAGTTTGACAT	1300
Dd		1448	ATCGTAATGAAGCAACGCATACCGGCGTGCAGTTTGTAAT	1487
RESULT 8				
LOCUS		AJ356572	1901 bp	DNA linear PAT 06-FEB-2002
DEFINITION		Sequence 7 from Patent WO0190333.		
ACCESSION		AJ356572		
VERSION		AJ356572.1 GI:18621059		
KEYWORDS		Escherichia coli.		
SOURCE		Escherichia coli.		
ORGANISM		Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;		
REFERENCE		Escherichia.		
AUTHORS		Short,J.M., Kretz,K.A., Gray,K.A., Barton,N.R., Garrett,J.B. and O'Donoghue,E.		
TITLE		Recombinant bacterial phylases and uses thereof		
JOURNAL		Patent: WO 0190333-A 7 29-NOV-2001;		
FEATURES		DIVERSA CORPORATION (US)		
SOURCE		location/Qualifiers 1..1901 /organism="Escherichia coli" /db_xref="taxon:562"		
BASE COUNT		474 A 499 C 499 G 428 T		1 others
ORIGIN				
Query Match		97.3%:	Pred. 1287.8;	DB 6; Length 1901;
Best Local Similarity		99.5%;	Freq. No. 0;	
Matches 1293; Conservative		0;	Mismatches 7;	Indels 0; Gaps 0;
OY		1	ATGAAGGATCTTAATCCCATTTTTTATCTCTCTGATTCCGTTAACCCGCAATCTGCA	60
Dd		188	ATGAAGGATCTTAATCCCATTTTTTATCTCTCTGATTCCGTTAACCCGCAATCTGCA	247
OY		61	TTGCGTCAGACTGACCGGAGCTGNAAGTGSAAAGTGTGTGATTGTCACTGCTCANGT	120
Dd		248	TTCCCTCAGAGTAGCCGCGAGCTGMAAGTGGAAAGTGTGTGATTGTCACTGCTCANGT	307
OY		121	GTTGGTGTCCAACAAGCCGACCAATGATGAGATGCAAGATGACCCCCAGACGATGGCA	180
Dd		308	GTTGGTGTCCAACAAGCCGACCAATGATGAGATGCAAGATGACCCCCAGACGATGGCA	367
OY		181	ACCTGGCCGATAAACTGGTTGGCTGACACCGCGNGTGGTGAATTCAGTCCATCTC	240
Dd		368	ACCTGGCCGATAAACTGGTTGGCTGACACCGCGNGTGGTGAATTCAGTCCATCTC	427
OY		241	GGACATTACCAAGCCACGCTGTGTGTAGCCGACGGATTCTGGCGAAAAAGGCTGCCG	300
Dd		428	GGACATTACCAAGCCACGCTGTGTGTAGCCGACGGATTCTGGCGAAAAAGGCTGCCG	487
OY		301	CAGTCTGGTCAAGTGGCGGATTAATGATGATTCGACGAGCGGTACCCGTAANAAGGGGAA	360
Dd		488	CAGTCTGGTCAAGTGGCGGATTAATGATGATTCGACGAGCGGTACCCGTAANAAGGGGAA	547
OY		361	GCTTTCCGCGCGGGGCGGACACTGTGTCAATTAACCGTACATACCCAGGACGATACG	420
Dd		548	GCTTTCCGCGCGGGGCGGACACTGTGTCAATTAACCGTACATACCCAGGACGATACG	607
OY		421	TCCAATCCCATCCGTTATTTTAATCTCTTA AAAA CTGGCGTTTGCCAACTGGATTAACGG	480
Dd		608	TCCAATCCCATCCGTTATTTTAATCTCTTA AAAA CTGGCGTTTGCCAACTGGATTAACGG	667
OY		481	AACGTGACTGACGGGATTCCTCAGAGAGGACAGGAGGTCAATTCGTGACTTTAACGGGGAT	540
Dd		668	AACGTGACTGACGGGATTCCTCAGAGAGGACAGGAGGTCAATTCGTGACTTTAACGGGGAT	727

```

QY 541 CGGCAACGGCGTTTCGCAACTGGAACGGGTCCTAATTTTCCGCAATCAACTGTGC 600
    |||||||
Db 728 CGGCAACGGCGTTTCGCAACTGGAACGGGTCCTAATTTTCCGCAATCAACTGTGC 787
QY 601 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGCATTACCAATCGAACTC 660
    |||||||
Db 788 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGAGCATTACCAATCGAACTC 847
QY 661 AAGGTGAGCGCGCAATGTCATTAACCGGTGCGGTAAAGCCATCAATCTGACG 720
    |||||||
Db 848 AAGGTGAGCGCGCAATGTCATTAACCGGTGCGGTAAAGCCATCAATCTGACG 907
QY 721 GAGATATTTCTCTGCAACAAGCAGCAGGAAATGCGGAGCGGGGAGAGATACC 780
    |||||||
Db 908 GAGATATTTCTCTGCAACAAGCAGCAGGAAATGCGGAGCGGGGAGAGATACC 967
QY 781 GATTACACACAGTGGAAACACCTTGCTAAGTTTGATACCGCAATTTTATTTGCTACAA 840
    |||||||
Db 968 GATTACACACAGTGGAAACACCTTGCTAAGTTTGATACCGCAATTTTATTTGCTACAA 1027
QY 841 CGGACGCGCAGAGTTGCGCGCAGCGCGCACCCCGTATTGATTTGATCATGCGACG 900
    |||||||
Db 1028 CGGACGCGCAGAGTTGCGCGCAGCGCGCACCCCGTATTGATTTGATCATGCGACG 1087
QY 901 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTACCCACTTCAGTACTG 960
    |||||||
Db 1088 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTACCCACTTCAGTACTG 1147
QY 961 TTTATTCGCGGACACGATCTAATCTGGCAATCTCGCGCGCAGCTGAGACTCACTG 1020
    |||||||
Db 1148 TTTATTCGCGGACACGATCTAATCTGGCAATCTCGCGCGCAGCTGAGACTCACTG 1207
QY 1021 ACGCTTCGCGGTAGCGCGGATACAGCGCGCAGGTGATCACTGCTTTGAAACGCTG 1080
    |||||||
Db 1208 ACGCTTCGCGGTAGCGCGGATACAGCGCGCAGGTGATCACTGCTTTGAAACGCTG 1267
QY 1081 CGTGCGGTAAGCGATTAACAGCAGTGTGATTCAGGTTTCGCTTCAGACTTTTACG 1140
    |||||||
Db 1268 CGTGCGGTAAGCGATTAACAGCAGTGTGATTCAGGTTTCGCTTCAGACTTTTACG 1327
QY 1141 CAGATGCGGTATTAACCGCGCGTCTCAATTAATACGCGCGCGGAGAGGTGAACG 1200
    |||||||
Db 1328 CAGATGCGGTATTAACCGCGCGTCTCAATTAATACGCGCGCGGAGAGGTGAACG 1387
QY 1201 CTGGCAGAGTGTGAAGCGCAAAATGGCAGGCGATGCTGCTGGAGAGTTTACGCA 1260
    |||||||
Db 1388 CTGGCAGAGTGTGAAGCGCAAAATGGCAGGCGATGCTGCTGGAGAGTTTACGCA 1447
QY 1261 ATCGTGAATGAGCAGCAGTACCGCGCTGCGAGTTTGAGAT 1300
    |||||||
Db 1448 ATCGTGAATGAGCAGCAGTACCGCGCTGCGAGTTTGAT 1487

```

```

RESULT 9
AE000200
LOCUS Escherichia coli K12 MG1655 section 90 of 400 of the complete
DEFINITION genome.
ACCESSION AE000200 U00096
VERSION AE000200.1 GI:2367111
KEYWORDS
SOURCE
ORGANISM Escherichia coli K12.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
AUTHORS Blatterer,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
Mau,B. and Shao,Y.
TITLE The complete genome sequence of Escherichia coli K-12
JOURNAL Science 277 (5311), 1453-1474 (1997)
MEDLINE 97426617

```

```

PUBMED 9278503
REFERENCE
AUTHORS Blatterer,F.R.
TITLE Direct Submission
JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
REFERENCE
AUTHORS Blatterer,F.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
Email: ecoligenetics.wisc.edu Phone: 608-262-2534 Fax:
608-263-7459
4 (bases 1 to 10784)
REFERENCE
AUTHORS Plunkett,G. III.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of
Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
COMMENT
On Sep 9, 1997 this sequence version replaced gi:1787214.
This sequence was determined by the E. coli Genome Project at the
University of Wisconsin-Madison (Frederick R. Blattner, director).
Supported by NIH grants HG00301 and HG01428 (from the Human Genome
Project and NCHGR). The entire sequence was independently
determined from E. coli K12 strain MG1655. Predicted open reading
frames were determined using Genemark software, kindly supplied by
Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
30332 (e-mail: markborodov.gatech.edu). Open reading frames that
have been correlated with genetic loci are being annotated with CG
site Nos., unique ID nos. for the genes in the E. coli Genetic
Stock Center (CGSC) database at Yale University, kindly supplied by
Mary Berlyn. A public version of the database is accessible
(http://cgsc.biology.yale.edu). Annotation of the genome is an
ongoing task whose goal is to make the genome sequence more useful
by correlating it with other data. Comments to the authors are
appreciated. Updated information will be available at the E. coli
Genome Project's World Wide Web site
(http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
its annotations are periodically updated; this is version M54. No
sequence changes. Annotation updates: updated gene identifications
and products; all new functional assignments courtesy of Monica
Riley; added promoters, protein binding sites, and repeated
sequences described in reference 1. The unique numeric identifiers
beginning with a lowercase 'b' assigned to each gene (protein- or
RNA-encoding) are now designated as gene synonyms instead of
labels. This should allow them to be searched for in Entrez as gene
names.
FEATURES
SOURCE location/Qualifiers
1..10784
/organism="Escherichia coli K12"
/strain="K12"
/sub_strain="MG1655"
/db_xref="taxon:83333"
49..84
/note="REP (repetitive extragenic palindromic) element;
contains 1 REP sequence"
repeat_region
gene
CDS

```

/translation="KAILIPPLSLILVTPQSAFQAQSEPEIKLESVIVYSHGVAP
 TKAQIMODVTPDAMPVPRKIGWLTTPRGGLIAVLGHYQORQLVADLLAKKQDPS
 GOVAIIVADDERKTGEAFAGLADCAIVYHQTADYSSPFLNPLKTCQCDNA
 NVTDAIISRAGSIAIDFTGHTQAFRELEVNFQSNICLKREKODSCSLTOLPS
 ELKYSADNVSLTGAASLASMLTEIFLQAQGMPEGRITDSQHWMTLSLHNAOF
 YLQRTPEVARSRAIPLDLIKTALTTPHPQKQATVLPVSLEFIAGHDITLALG
 ALELWMLPGQPDNTPGGELVFERMRRLSDNSQMIQVSLVFQTLQMKMDKPTLSLT
 PPGEVKTLPGCEENRAGMCSLAFQIVNRAIPACSL"
 1153. .1189
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 repeat_region
 1491.1524
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 gene
 complement(1536. .3716)
 /gene="yccc"
 /note="b0981"
 complement(1536. .3716)
 CDS
 complement(1536. .3716)
 /gene="yccc"
 /function="orf; unknown"
 /note="f726; 100 pct identical to fragment YCCC_ECOLI
 SM:P38134 but has 627 additional N-terminal aa and 2 less
 C-terminal residues"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical protein"
 /protein_id="AAC74066.1"
 /db_xref="GI:187216"
 /translation="MTKNMTPGSGTOENETDLRLVGLMDHRRFIISVATFLI
 AVASLSTPIYQADTIVQVQKQCNALISGSDMIPSSPESAPETIOLLSRMILK
 TIAELNLDIVQKFTPIVGRMARITKREBELISMHILPOLNGQDQDLTVGEN
 GHYTELEGEFTVNGVQRLERKQVALITADIKAPGTFVLSQRTLEALNALDET
 TVSERKSGMLKELMTGDDPOLITRIINLSANNYIQAQADSDSLEIFQRL
 PEVSELDQAQEKLVYRQQRSDVLEAKAVLEQIVNVDQLNELTFRELEISOL
 KQDHTFALLERKOTLEOERKRLNKRVASMSQOEVLIRSDVAGAVYARMLR
 QDELISKSASIAIGNRIIDPAVTOPOPVKPKALNVVGLTGLFISGAVYARMLR
 RGVAPRQLEHGISVYATIPMSSEWLDKTRIRKKNLSNQRHRTKINPIAVNPA
 DSAVAVALKRSLHFAMETENNILMTGTAPDSGKTFVSTLAIVAQSQKVLFI
 DADLRGYSNHLFTVSNHGLSEYLAGDELKRYIQHKGGEFDVITRGQVPPNSL
 LMRDMLQELMANDHYLDVDPMPMLAVSPAAYGVRSGVSLVAARFGLTAKESV
 LSMORLEQAGVNIKGAIDNGVYIKRASTAVSYGVYYSSEKE"
 complement(3733. .3764)
 /note="factor Sigma70; predicted +1 start at 1043443"
 promoter
 complement(3736. .4194)
 /gene="yccy"
 /note="b0982"
 complement(3736. .4194)
 CDS
 complement(3736. .4194)
 /gene="yccy"
 /note="b0982"
 complement(3736. .4194)
 /function="putative enzyme; Not classified"
 /note="f152; This 152 aa ORF is 33 pct identical (13 gaps)
 to 139 residues of an approx. 160 aa protein PPAI_SCHPO
 SM:P41893"
 /codon_start=1
 /transl_table=11
 /product="putative phosphatase"
 /protein_id="AAC74067.1"
 /db_xref="GI:1787217"
 /translation="MDSNAQLKFNLSILVCTGNICRSPIGERLIRKRLPGYKVSAG
 VHGVKHPADPADVADVAVNAGVLEGHAGARKLTAMARNYDIIILAMESENIAOVYATA
 PEVGRKTMLEFGOMLEQKEIPDPYRKSQDAFHYVGMLEKRASEMARRLSR"
 complement(3793. .3820)
 promoter
 complement(3793. .3820)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043503"
 promoter
 complement(3818. .3844)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043528"
 promoter
 complement(3865. .3892)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043575"
 gene
 complement(4170. .5309)
 /gene="ycc2"
 /note="b0983"
 complement(4170. .5309)
 CDS
 complement(4170. .5309)

/gene="ycc2"
 /function="phenotype; Not classified"
 /note="f379; This 379 aa ORF is 29 pct identical (33 gaps)
 to 364 residues of an approx. 400 aa protein BEXD_HMEIN
 SM:P22236"
 /codon_start=1
 /transl_table=11
 /product="putative function in exopolysaccharide
 production"
 /protein_id="AAC74068.1"
 /db_xref="GI:187218"
 /translation="MKKNIFKRSVLTAVLSTACTVPGQNLSTKNQDVELPNOY
 DIDKNVITPVTEGLIDIDIRAPRIMSQAPELEQDIANTERIGIDVLMVWHPPE
 LTPAGQYRSASADTGNWVNDALIFYPYIGRLKAGKLTQVRNEITATLADSVSPQ
 VDVSAFRSOKAYVTGEYSKGOQPTINIPITINDALNAAGLTADADMRNVLTQ
 GVTQVNYIALMQRGDLRQNLKHPDGLFIRNDLKFVWGEVGEKOSTLMDRSGM
 TLAEALGNAEGNODVADATGTEFVIRATQONKGTANTYOLNARDASMLIGTEFOL
 EPTDIVYVYTRAPLARWNRYSILVPTISGVHDLTETSRNIQTWPN"
 complement(5355. .7451)
 gene
 complement(5355. .7451)
 /gene="ymca"
 /note="b0984"
 complement(5355. .7451)
 CDS
 complement(5355. .7451)
 /gene="ymca"
 /function="orf; unknown"
 /note="f698; This 698 aa ORF is 64 pct identical (7 gaps)
 to 694 residues of an approx. 704 aa protein yJRH_ECOLI
 SM:P32689"
 /codon_start=1
 /transl_table=11
 /product="orf, hypothetical protein"
 /protein_id="AAC74069.1"
 /db_xref="GI:187219"
 /translation="MKKSQYLSCIAAVSSNAEVLTPDPGLSSQSDPFGTGLQ
 MNARIAPEGSESVNYRNDQYRFSTSVLALPMLEGTIRYDVRTRYQSOWEDSGD
 OSYKXSPDFKRLMEGFWLQVAVAFGRKDIGTGLFQDEYVIAKQAGPFEFTGMA
 KQYANACNITNTPCRVSDKXCHRAHSDAGDISDIFRGASIFGIEFYOTPNPL

Query Match 97.3%; Score 1287.8; DB 1; Length 10784;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;

1 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCCGAATCTGA 60
 123 ATGAAGCGATCTTAATCCATTTTATCTCTGATTCGCTTAACCCCGAATCTGA 182
 61 TTCGCTCAGAGTACGCGGAGCTGAAGCTGAAAGTGTGATTCGCTAGTGTATG 120
 183 TTCGCTCAGAGTACGCGGAGCTGAAGCTGAAAGTGTGATTCGCTAGTGTATG 242
 121 GTGCGTCTCAACCAAGGCCACGACATGATGACAGATGTACCCAGACCATGGCA 180
 243 GTGCGTCTCAACCAAGGCCACGACATGATGACAGATGTACCCAGACCATGGCA 302
 181 ACCGTGGCGGTAAACTGGGTTGGCTGACACCGCGGNGGGAGCTATATGCCATATC 240
 303 ACCGTGGCGGTAAACTGGGTTGGCTGACACCGCGGNGGGAGCTATATGCCATATC 362
 241 GCACATTACCAAGCGCAGCGTGTGTAGCGGAGATTGCTGGCAAAAAGGGCTGCCG 300
 363 GCACATTACCAAGCGCAGCGTGTGTAGCGGAGATTGCTGGCAAAAAGGGCTGCCG 422
 301 CAGTGTGTGACGCTGCGATTATTCGTGATGTGACGACGACGATACCCGTAAACAGCGCA 360
 423 CAGTGTGTGACGCTGCGATTATTCGTGATGTGACGACGACGATACCCGTAAACAGCGCA 482
 361 GCCTTGCGCGCGCGGCTGGCACCTGACGTGCAATAACGCTATACCCAGGACAGATCG 420
 483 GCCTTGCGCGCGCGGCTGGCACCTGACGTGCAATAACGCTATACCCAGGACAGATCG 542
 421 TCACATCCCGATCCCTTATTTATCTCTAATAAAGCTGGCGTTGGCAACTGGATAACGCG 480
 543 TCACATCCCGATCCCTTATTTATCTCTAATAAAGCTGGCGTTGGCAACTGGATAACGCG 602

```

QY 481 AACGTCAGTACCGATCTCTACAGAGGACGAGAGGTCAATTCGCTACTTACCGGCAT 540
    |||
Db 603 AAGTGTACTACCGATCTCTACAGAGGACGAGAGGTCAATTCGCTACTTACCGGCAT 662
QY 541 CGGCAAAAGCGGTTTGGCGAATGGAAGGGTGTATATTTTCGCAATCAAACTTGTG 600
    |||
Db 663 CGGCAAAAGCGGTTTGGCGAATGGAAGGGTGTATATTTTCGCAATCAAACTTGTG 722
QY 601 CTTAAACGTGAGAAAGGACGAAAGGTCTTCAATTAACGAGCATTAACATCGGAATC 660
    |||
Db 723 CTTAAACGTGAGAAAGGACGAAAGGTCTTCAATTAACGAGCATTAACATCGGAATC 782
QY 661 AAGGTAGCGCGCAGCAATCTCTAATTAACGAGGTGCGGTAAAGCTTCGATCAATCTGACG 720
    |||
Db 783 AAGGTAGCGCGCAGCAATCTCTAATTAACGAGGTGCGGTAAAGCTTCGATCAATCTGACG 842
QY 721 GAGATATTTCTCTGCAACAAGCAAGCAAGGAATGCCGAGCGGGGGTGGGAAGGATCAC 780
    |||
Db 843 GAGATATTTCTCTGCAACAAGCAAGCAAGGAATGCCGAGCGGGGGTGGGAAGGATCAC 902
QY 781 GATTCACACAGTGAAGAACACCTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 840
    |||
Db 903 GATTCACACAGTGAAGAACACCTGCTAAGTTTGATTAACGCGCAATTTATTTGCTACAA 962
QY 841 CGCAGCGCAGAGGTTGGCCGCGCAGCGCGCACCCGTTATTTGATTTGATCATGCGACG 900
    |||
Db 963 CGCAGCGCAGAGGTTGGCCGCGCAGCGCGCACCCGTTATTTGATTTGATCATGCGACG 1022
QY 901 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACG 960
    |||
Db 1023 TTGACGCGCCATCCACCGCAAAAACAGCGGTATGCTGACATTAACCACTTCAGTACG 1082
QY 961 TTATATGCGGAGACAGTACTAATCTGGAATCTGCGCGCGCAGCGAGGCTCAACTG 1020
    |||
Db 1083 TTATATGCGGAGACAGTACTAATCTGGAATCTGCGCGCGCAGCGAGGCTCAACTG 1142
QY 1021 ACGCTTCCCGGTAGCGCGGATTAACACCGCGCAGGTGTAACCTGCTTTGAACGCTG 1080
    |||
Db 1143 ACGCTTCCCGGTAGCGCGGATTAACACCGCGCAGGTGTAACCTGCTTTGAACGCTG 1202
QY 1081 CGTCGCGTAAAGCGATTAACAGCGAGGTGATTCAGTTTCGCTTCACAGCTTTACG 1140
    |||
Db 1203 CGTCGCGTAAAGCGATTAACAGCGAGGTGATTCAGTTTCGCTTCACAGCTTTACG 1262
QY 1141 CAGATGCGTATAAAAGCGCGGTCTCAATTAATACGCGCGCGGAGAGGTGAACCTGAC 1200
    |||
Db 1263 CAGATGCGTATAAAAGCGCGGTCTCAATTAATACGCGCGCGGAGAGGTGAACCTGAC 1322
QY 1201 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGAGCATGTTGCTGGCAGGTTTACGCA 1260
    |||
Db 1323 CTGGCAGGATGTGAAGAGCAAAATGCGCAGGAGCATGTTGCTGGCAGGTTTACGCA 1382
QY 1261 ATCGTGAATGAAGCAGCATACCGGCTGAGTTTGAGAT 1300
    |||
Db 1383 ATCGTGAATGAAGCAGCATACCGGCTGAGTTTGAGAT 1422

RESULT 10
D90735 18882 bp DNA linear BCT 07-FEB-1999
LOCUS Escherichia coli genomic DNA (22.3 - 22.7 mln).
DEFINITION D90735 AB001340
ACCESSION D90735.1 GI:4062539
KEYWORDS Complete and shotgun sequencing; hyaB; hyaC; hyaD; hyaE; hyaF;
appC; cyxA; cysB; yccM; appB; cyxB; cysD; appA; amsI; insA; insB; cspF;
cspB; sfa; yccM; yccM.
SOURCE Escherichia coli (strain:K12) DNA, clone:Kohara clone #225.
ORGANISM Escherichia coli
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (sites)
AUTHORS Oshima,T., Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
Ikemoto,K., Inada,T., Itoh,T., Kajihara,M., Kanai,K., Kashimoto,K.,

```

```

TITLE JOURNAL
MEDLINE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
COMMENT

Kimura,S., Kitagawa,M., Makino,K., Masuda,S., Miki,T.,
Mizobuchi,K., Mori,H., Motomura,K., Nakamura,Y., Nishimoto,H.,
Nishio,Y., Saito,M., Sampel,G., Seki,Y., Tagami,H., Takemoto,K.,
Wada,C., Yamamoto,K., Yano,M., and Horiiuchi,T.
A 718-kb DNA sequence of the Escherichia coli K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map
DNA Res. 3 (3), 137-155 (1996)
97061202
2 (sites)
Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A., Horiiuchi,T.,
Ikemoto,K., Inada,T., Isono,K., Isono,S., Itoh,T., Kanai,K.,
Kasai,H., Kashimoto,K., Kim,S., Kimura,S., Kitagawa,M.,
Kitakawa,M., Makino,K., Masuda,S., Miki,T., Mizobuchi,K., Mori,H.,
Motomura,K., Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T.,
Saito,M., Sampel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
Yamamoto,Y., and Yano,M.
The systematic sequencing of the Escherichia coli genome in Japan
Unpublished
3 (bases 1 to 18882)
Direct Submission
Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science
and Technology, Res. & Edu. Center for Genetic Info.; 8916-5
Takayama, Ikoma, Nara 630-01, Japan
(E-mail:hmori@igc.aist-nara.ac.jp, Tel:81-7437-2-5660,
Fax:81-7437-2-5669)
On Dec 24, 1998 this sequence version replaced gi:1651473.
Collaboration Information:
Project:
The Japan E.coli genome DNA sequencing project
Group:
The Japan E.coli genome DNA sequencing group
Members: (1995.4 - 1996.3)
Alba,H., Baba,T., Fujita,K., Hayashi,K., Honjo,A.,
Horiiuchi,T., Ikemoto,K., Inada,T., Isono,K., Isono,S.,
Itoh,T., Kanai,K., Kasai,H., Kashimoto,K., Kim,S.,
Kimura,S., Kitagawa,M., Kitakawa,M., Makino,K.,
Masuda,S., Miki,T., Mizobuchi,K., Mori,H., Motomura,K.,
Nakamura,Y., Nishimoto,H., Nishio,Y., Oshima,T., Saito,M.,
Sampel,G., Seki,Y., Tagami,H., Takemoto,K., Wada,C.,
Yamamoto,Y., and Yano,M.
Headed by:
Name: Takashi Horiiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishiorien@nibb.ac.jp
Information operator:
Name: Hirotsada Mori
Address: NARA Institute of Science and Technology,
Ikoma, 630-01, Japan
E-mail: hmori@igc.aist-nara.ac.jp
URL:
The Japan E. coli genome database
http://sw3.aist-nara.ac.jp.
Location/Qualifiers
1..18882
/organism="Escherichia coli"
/strain="K12"
/db_xref="taxon:562"
/map="22.3-22.7 mln"
/clone="Kohara clone #225"
/note="Nucleotide position 1034630-1053511 from the
initiation site of Phra (0 min). This clone is from
Kohara lambda miniset library."
1..966
/gene="hyaB"
<1..966
/gene="hyaB"
/note="ORF_ID:0224#6
similar to SwissProt Accession Number P19927"
/codon_start=1
/transl_table=1
/product="Hydrogenase-1 large chain (EC 1.18.99.1) (nife
hydrogenase) (membrane-bound hydrogenase 1 large subunit)

```

```

(hyd1)."
/protein_id="BAA35738.1"
/db_xref="GI:4062540"
/translation="PDALAIGOFNKPMSEIGTGLSDKCVLSYGAPDIAINDGEKSL
MPGAVINGDNRNNVLPDVLVDQVDFEYVADHMYRYPNDQYGRHPFDGTTDWTWFGD
VKGSDNTIQOLNEDEERYSWIKAPRWGRNAMEGWLPARTLIAVHKGDATVSEVDMMS
ALNLPISGISTQILICRAHAEQMAACKLOYPFDKMLTNTLKNGLMATATEKEMPAT
WPTECRGVGTETAPRGALGHMAAIKIDGKIDLYOCVYPTTWNASPRDPKQIGAYEBAAL
MNTKMAIPEOPLELRLTSHSFPCLACSTHYLGDSGLISVQR"
85. 1692
/gene="hyac"
/contig="hyac"
/feature="hyac"
/feature_start="1"
/feature_end="1692"
/feature_note="ORF_ID:0225#1
similar to PIR Accession Number JY0074"
/translation="MOOKSDNVSHYVEAPVRIWMLTVLCMAVLMTGYEIGKPLP
SVSEATLYFYNGYRIHLHFSAWGFVYVLLMRYIAVGNRYSRELFIVPMKRSW
OGVWEIRWYLFELKRPASDIGHNPIAQAMPGYLMAFGYMLITGELFALYSOVAIF
APRYVEFEYFTGNSMDIHSWRLGMLIGAFVIGHYVALREDIMSDTVISTMV
NGIRSHKRGKISNERS"
1689. 2276
/gene="hyad"
/contig="hyad"
/feature="hyad"
/feature_start="1"
/feature_end="2276"
/feature_note="ORF_ID:0225#2
similar to SwissProt Accession Number P19930"
/translation="Hydrogenase-1 operon protein hyad."
/protein_id="BAA35740.1"
/db_xref="GI:1651476"
/translation="MSEORVYVMGLGNLIMADDEGCVAVARLTAHYHMPVEYELVDG
GTGILNGLYVESASHLIDAIDGLDEGGLRTYAGERIPAYISAKKMSLQNSFSE
VIALADINGHLPAHIALVGLDPAIMLDGEGSLSELARQPLPAEAQALAOALAMGIVP
QPANESRLNDCSLMENYEGVRLRQYRMTOEQG"
2273. 2671
/gene="hyae"
/contig="hyae"
/feature="hyae"
/feature_start="1"
/feature_end="2671"
/feature_note="ORF_ID:0225#3
similar to SwissProt Accession Number P19931"
/translation="Hydrogenase-1 operon protein hyae."
/protein_id="BAA35741.1"
/db_xref="GI:1651477"
/translation="MSNDTPDALMQRLMANGWTPVSESRILDMITQAPDGVLLSSD
PKRPPEVSDNPMVIGELREPPDYTMQVADLRQSEALIGDRFGVFRPATLVFTGN
YRGLNGIHMPAELINLMRGLVEPQOEAS"
2668. 3525
/gene="hyaf"
/contig="hyaf"
/feature="hyaf"
/feature_start="1"
/feature_end="3525"
/feature_note="ORF_ID:0225#4
similar to SwissProt Accession Number P19932"
/translation="Hydrogenase-1 operon protein hyaf."
/protein_id="BAA35742.1"
/db_xref="GI:1651478"
/translation="MSEFPHLIGRGTOPNDSSMNPITPCOVNDEPSMAALEQCA
HSRQVIALNELQHLSEKRPPLGEVYLVADLNLNADNRHTINTILGEGESVRIQDA
DSESEIOEALIFCGLMWRARRRGEKLEKLEAGCAPALAOAONLPLPDSLLPP
IDGLMNGILPLAHELIAHYRNDQAPHSINDKLPISEADRLFLSLCGPNIQIITIG
YGESYINATGLRHVWHLRCDTLKGPLLESYEICPIPEVLAAPEDLVDSARLSEVC
QMLAEAPT"
3659. 5203
/gene="appc"
/contig="appc"

```

```

CDS
/contig="cdda: cyxa"
/feature="appc"
/feature_start="1"
/feature_end="5203"
/feature_note="ORF_ID:0225#5
similar to SwissProt Accession Number P26459"
/translation="Cytochrome bd-11 oxidase subunit 1 (EC
1.10.3.-)."
/protein_id="BAA35743.1"
/db_xref="GI:1651479"
/translation="MWDVIDSRMOPALTYHLEFVPLTGLIFLAIMETIYVYG
RTIYRDMTRFWGKLEFGINFALGVATGLTMEFQFNSFYSNVTGDIAGPLAMEALM
AEFLSTFVGLFEFGMRQRLNKHQHLVWLVAFGSNLSALMILNANGMWOYPTGAHFD
IDTLREMTSESELVFNPSQVKEVYTAGVYGAEMTAMISAMYLGRGERNVALLR
SPAIGSVFTGLAIICTLOGDSASAYEAVOVKTAAMEGQTEPAPAFVAVAPE
QDOERNAPALKRIPALGILATHSIDKPVGKRNIAEMTYPRLOGRMAMLMQETISOG
NREPHVIOAFRLEBDLGTGMLISRYAPDNHRYTAQYOQAMRGAITPOVAPYNSERI
MVGCSLLVWLIALVOTLRKIDQHMVLMALMSLPLPAIAIEAGMFTERGKRP
MAIDILPTYSNHSALTGTGOLAFSLIMIVGLYTFLLIAEYVLMOKYARLGPANQSEQ
PTQOOG"
5215. 6351
/gene="appb"
/contig="appb"
/feature="appb"
/feature_start="1"
/feature_end="6351"
/feature_note="ORF_ID:0225#6
similar to SwissProt Accession Number P26458"
/translation="Cytochrome bd-11 oxidase subunit 11 (EC
1.10.3.-)."
/protein_id="BAA35744.1"
/db_xref="GI:1651480"
/translation="MEDYETLRFITWMLIGVILVYFMISDGFDMIGCLLPVARNDD
ERRIYNSVGAHMEBNQVWMLIAGALTAANPRVYAAAFSGFYAMITVLSIFRPL
AFDYRGTADAKRMWDAGLVISGVPPVFGINFGNLLGVPAPFQRLREVLGS
FMQWLTPPPLCGLLSLGMNVILQGGVWLQLTAVGVIHLRSQLATRAALLVWLCTFLA
GYLWLVGIDGVPVLLAQDANGSPNPLMKLVAVYLPGMANNFVSEPVLMIFPLGFCPL
LTVMAIYGRGRGWFMLASLMQGFVITAFYITGLTFPFWMPSSVPSISLTLMDSSTSQL
Query Match 97.3%; Score 1287.8; DB 1; Length 18882;
Best Local Similarity 99.4%; Pred. No. 0;
Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
1 ATGAAGCGACTATATCCATTTTATCTCTTGATTCCTTAACCCGCAATCTGCA 60
6536 ATGAAGCGACTATATCCATTTTATCTCTTGATTCCTTAACCCGCAATCTGCA 6595
61 TTGCTCAGAGTAGACCGGAGCTGAAGCTGGAAGTGTGGTATTGTCAATCGTCATG 120
6596 TTGCTCAGAGTAGACCGGAGCTGAAGCTGGAAGTGTGGTATTGTCAATCGTCATG 6655
121 GTGCGTGTCCAAACCAAGCCACGACGACGATGAGATGTCAGCCGACGACGATGGCA 180
6656 GTGCGTGTCCAAACCAAGCCACGACGACGATGAGATGTCAGCCGACGACGATGGCA 6715
181 ACCTGCGCGTAAACTGGGTGCTGACACCGCGGNGTGTGATGACTATGCTATCTC 240
6716 ACCTGCGCGTAAACTGGGTGCTGACACCGCGGNGTGTGATGACTATGCTATCTC 6775
241 GGACATTACCAAGCCGACGCTGTGTAGCCGACGATGCTGCGGCAAAAAGGCTGCCG 300
6776 GGACATTACCAAGCCGACGCTGTGTAGCCGACGATGCTGCGGCAAAAAGGCTGCCG 6835
301 CAGTGTGCTCAGTGTGCGCATTTATCTGATGCGACGAGCGATGACCCGTAACACAGCGAA 360
6836 CAGTGTGCTCAGTGTGCGCATTTATCTGATGCGACGAGCGATGACCCGTAACACAGCGAA 6895
361 GCCTTGGCGCGCGGCTGACACCTGACTGTGCAATAACCGTACATACCCAGCAGATACG 420
6896 GCCTTGGCGCGCGGCTGACACCTGACTGTGCAATAACCGTACATACCCAGCAGATACG 6955

```

QY 421 TCAGTCCGACGCGTATTATTCCTCTTAAACATGCGCTTTCCTCAACTGATACGCG 480
 DB 6956 TCAGTCCGACGCGTATTATTCCTCTTAAACATGCGCTTTCCTCAACTGATACGCG 7015
 QY 481 AACGTGACTGACGGATCTCTCAGACGAGGAGGCTCAATTTGCTGACTTTACGGGCAT 540
 DB 7016 AACGTGACTGACGGATCTCTCAGACGAGGAGGCTCAATTTGCTGACTTTACGGGCAT 7075
 QY 541 CGGCAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 600
 DB 7076 CGGCAACGGCGTTTCGCAACTGGAACGGGTGCTTAATTTTCGCAATCAAACTTGTC 7135
 QY 601 CTTAAACGTGAGAAACAGAGCAAGAGTGTTCATTAAACGAGCATTAACATCCGAATC 660
 DB 7136 CTTAAACGTGAGAAACAGAGCAAGAGTGTTCATTAAACGAGCATTAACATCCGAATC 7195
 QY 661 AAGGTGAGCGCCGCAATGTCTCATTTAACCGGTGCGGTAGCCTGCAATCAATGCTGACG 720
 DB 7196 AAGGTGAGCGCCGCAATGTCTCATTTAACCGGTGCGGTAGCCTGCAATCAATGCTGACG 7255
 QY 721 GAGATATTTCTCCGCAACAAAGCAGAGGAATGCGGAGCGGGGTGGGGAAGATCAC 780
 DB 7256 GAGATATTTCTCCGCAACAAAGCAGAGGAATGCGGAGCGGGGTGGGGAAGATCAC 7315
 QY 781 GATTCAACACAGTGAACACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTGCTACAA 840
 DB 7316 GATTCAACACAGTGAACACACCTTGTCTAAGTTTGATTAACGCGCAATTTTATTGCTACAA 7375
 QY 841 CGCACGCGAAGGTTGCCCGACGCGGCAACCCCGTTATTGATTCATCAGTGCACG 900
 DB 7376 CGCACGCGAAGGTTGCCCGACGCGGCAACCCCGTTATTGATTCATCAGTGCACG 7435
 QY 901 TTGACGCGCCATCCACGCGCAAAACAGGCGATGCTGATACCATTAACCATCTTCACTG 960
 DB 7436 TTGACGCGCCATCCACGCGCAAAACAGGCGATGCTGATACCATTAACCATCTTCACTG 7495
 QY 961 TTTATTCGCGGACAGATTAATCTGCAAAATCTCGGCGGCGCACTGAGCTCAACTG 1020
 DB 7496 TTTATTCGCGGACAGATTAATCTGCAAAATCTCGGCGGCGCACTGAGCTCAACTG 7555
 QY 1021 ACGCTTCCGCGTACGCGGATACACGCGCCGACGCTGATGTAATGCTGTTGAACGCTG 1080
 DB 7556 ACGCTTCCGCGTACGCGGATACACGCGCCGACGCTGATGTAATGCTGTTGAACGCTG 7615
 QY 1081 CGTCGGGTAAAGCGATTAACAGCAAGTGAATTCAGGTTTCGCTGCTTCCAGACTTACG 1140
 DB 7616 CGTCGGGTAAAGCGATTAACAGCAAGTGAATTCAGGTTTCGCTGCTTCCAGACTTACG 7675
 QY 1141 CAGATGCGGTATAAAGCGCGCTGTCAATTAATACGCGCGCCGAGAGAGTAAATGAC 1200
 DB 7676 CAGATGCGGTATAAAGCGCGCTGTCAATTAATACGCGCGCCGAGAGAGTAAATGAC 7735
 QY 1201 CTGCGAGATGTGAAGAGCAAAATGCGAGGCGATGTTCGTTGGCAGAGTTTAAACG 1260
 DB 7736 CTGCGAGATGTGAAGAGCAAAATGCGAGGCGATGTTCGTTGGCAGAGTTTAAACG 7795
 QY 1261 ATGCTGAATGAAGACCATTAACGCGGCTGCAAGTTGAGAT 1300
 DB 7796 ATGCTGAATGAAGACCATTAACGCGGCTGCAAGTTGAGAT 7835

RESULT 11

ECOCAPHB 1901 bp DNA linear BCT 04-Apr-2002
 LOCUS
 DEFINITION Escherichia coli periplasmic phosphoanhydride phosphohydrolase
 (appa) gene, complete cds.
 ACCESSION L03371.1 GI:145178
 VERSION L03371.1
 KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
 SOURCE Escherichia coli.
 ORGANISM Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.

REFERENCE 1 (sites)
 AUTHORS Dassa, J., March, C. and Boquet, P. L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appa reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616
 PUBMED 2168385
 REFERENCE 2 (bases 1 to 1901)
 AUTHORS Ostanin, K., Harns, E. H., Stevis, P. E., Kuciel, R., Zhou, M. M. and Van Ertten, R. L.
 TITLE Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase
 JOURNAL J. Biol. Chem. 267 (32), 22830-22836 (1992)
 PUBMED 93054596
 FEATRES 1429631
 source Location/Qualifiers
 1. 1901
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 188. 1486
 /gene="appa"
 188. 1486
 /standard_name="acid phosphatase"
 /gene="appa"
 /codon_start=1
 /translation="MKAILIPFLSLIPPLPPOSAPSOSEPELKLESVIVSRNGBRAP
 TKAQIMQDTPDAMPMPYKVLGILTRGSELIVLGHYQORLYANGLAKKCPROS
 GQVALIDVDNRTKRGARAGAPDCAITVHQADTSFDPLEPLKTEGCOLDA
 NVTADILISRAGSIADEFGHROTAPELELYNTPQSLCKRKOEDSCSLTGLDS
 ELKASDNVSLTGAVSLASMLTEIFLQAOAGPEPGNRTDSDHONVTLISNAGF
 YLQRTPEVARSRAPIPLDLKTLATLPHPPQKQAGVTLPSVFLIGHDNILNIG
 ALLENVTLPGQPDNTPPGSELVEFRWRRLSPNSIOVSIYVFTLQGMKDXPLSLMT
 PRGEVKLTLAGCEERNMAGMSLNGFQIVVNAIRPACSL"
 302. 304
 /gene="appa"
 /citation=[1]
 /phenotype="H17N mutant"
 BASE COUNT 475 a 498 c 499 g 428 t 1 others
 ORIGIN
 Query Match 97.2%; Score 1286.2; DB 1; Length 1901;
 Best Local Similarity 99.4%; Pred. No. 0;
 Matches 1292; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
 QY 1 ATGAAGCGATCTTAATCCCATTTTATCTTTCGATTCGCTTAACCCCGCAATGCA 60
 DB 188 ATGAAGCGATCTTAATCCCATTTTATCTTTCGATTCGCTTAACCCCGCAATGCA 247
 QY 61 TTCGCTGAGTGAAGCGCGAGCTGAAGCTGGAAGTGGTGGATTCAGTGCATAGGT 120
 DB 248 TTCGCTGAGTGAAGCGCGAGCTGAAGCTGGAAGTGGTGGATTCAGTGCATAGGT 307
 QY 121 GTGCGTGTCCCAACCAAGCGCACCAACTGATGAGATGTACACCCAGACGATGGCA 180
 DB 308 GTGCGTGTCCCAACCAAGCGCACCAACTGATGAGATGTACACCCAGACGATGGCA 367
 QY 181 ACCTGGCGGTAAACAGTGGTGGCTGACACCGGNGTGGTGGATTAATGCGCTATGTC 240
 DB 368 ACCTGGCGGTAAACAGTGGTGGCTGACACCGGNGTGGTGGATTAATGCGCTATGTC 427
 QY 241 GGACATTACCAACGCGACGCTGTGTAGCCGACGATGCTGGCGAAAAAGGCTGCCG 300
 DB 428 GGACATTACCAACGCGACGCTGTGTAGCCGACGATGCTGGCGAAAAAGGCTGCCG 487
 QY 301 CAGTCTGTGTCAGTGGCGATTAATGCTGATGTCAGACGAGCTACCCGTAACAGCGCA 360
 DB 488 CAGTCTGTGTCAGTGGCGATTAATGCTGATGTCAGACGAGCTACCCGTAACAGCGCA 547

OY	781	TAATTCACACCCAGTGGAAACACCTTCTCTAAGTTTGGATTAACGGCGAAATTTATTTCTCTCAAA	840
Db	968	GAATTCACACCCAGTGGAAACACCTTCTCTAAGTTTGGATTAACGGCGAAATTTATTTCTCTCAAA	1027
OY	841	CGCACGGCCAGAGGTGGCCCGCAGCCGGCCACCCCGCTTATTTGATTTGATCTATGTGCGACGG	900
Db	1028	CGCACGGCCAGAGGTGGCCCGCAGCCGGCCACCCCGCTTATTTGATTTGATCTATGTGCGACGG	1087
OY	901	TTTGACGGCCCATTCACCGGCAAAAACAGGCGATATGTTGACATTATACGCACATTTCAGTACTG	960
Db	1088	TTTGACGGCCCATTCACCGGCAAAAACAGGCGATATGTTGACATTATACGCACATTTCAGTACTG	1147
OY	961	TTTATTTGCGCGGACACGACTACTTAATCTGGCAATCTCGGCGGCGCAGTGGAGACTCAATCG	1020
Db	1148	TTTATTTGCGCGGACACGACTACTTAATCTGGCAATCTCGGCGGCGCAGTGGAGACTCAATCG	1207
OY	1021	ACGCTTCCCGGTACAGCCGGATTAACAGCCCGCCAGGTGGTGAATCTGTGTTGAACGCTGG	1080
Db	1208	ACGCTTCCCGGTACAGCCGGATTAACAGCCCGCCAGGTGGTGAATCTGTGTTGAACGCTGG	1267
OY	1081	CGTGGCGCTTAACGCAGTAACAGCCAGCTGTGATTCAGGCTTTCGCGTCTTCACAGCTTTTACAG	1140
Db	1268	CGTGGCGCTTAACGCAGTAACAGCCAGCTGTGATTCAGGCTTTCGCGTCTTCACAGCTTTTACAG	1327
OY	1141	CAGATGCGTGAATAAAACGCCGCTCTCATTTAATACGCCGCCGAGAGAGGTGAATCTGACC	1200
Db	1328	CAGATGCGTGAATAAAACGCCGCTCTCATTTAATACGCCGCCGAGAGAGGTGAATCTGACC	1387
OY	1201	CTGGCAGAGTATGAAGAAGCAAAATGCGCAGAGGCAATGTTCGTTGGCAGGTTTACGGAA	1260
Db	1388	CTGGCAGAGTATGAAGAAGCAAAATGCGCAGAGGCAATGTTCGTTGGCAGGTTTACGGAA	1447
OY	1261	ATCGTGAATGAAGCAACGCATACCGGCGGTGCACGTTTGAAGT	1300
Db	1448	ATCGTGAATGAAGCAACGCATACCGGCGGTGCACGTTTGAAGT	1487

LOCUS	AX356574	1901 bp	DNA	Linear	PAT 06-FEB-2002
DEFINITION	Sequence 9 from Patent WO0190333.				
ACCESSION	AX356574				
VERSION	AX356574.1	GI:18621060			
KEYWORDS					
SOURCE					
ORGANISM	Escherichia coli.				
	Escherichia coli				
	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.				
REFERENCE	1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.				
AUTHORS	Recombinant bacterial phylases and uses thereof				
TITLE	Patent: WO 0190333-A 9 29-NOV-2001;				
JOURNAL	DIVERSA CORPORATION (US)				
FEATURES	Location/Qualifiers				
source	1..1901				
	/organism="Escherichia coli"				
	/db_xref="taxon:562"				
BASE COUNT	473 a	499 c	500 g	428 t	1 others
ORIGIN					
Query Match	97.2%; Score 1286.2; DB 6; Length 1901;				
Best Local Similarity	99.4%; Pred. No. 0;				
Matches 1292:	Conservative	0;	Mismatches	8;	Indels 0; Gaps 0;
0Y	1 ATGAAGGATCTTAATCCATTTTATATCTTGTGATTCCGTTAACCCGCAATCTGCA 60				
Db	188 ATGAAGGATCTTAATCCATTTTATATCTTGTGATTCCGTTAACCCGCAATCTGCA 247				
0Y	61 TTCCCTGAGATGAGCCGGAGCTGAAAGTGTGGATGTTCAGTCGTGATGCT 120				
Db	248 TTCCCTGAGATGAGCCGGAGCTGAAAGTGTGGATGTTCAGTCGTGATGCT 307				

QY	121	GTGCGTGTCCAAACCAAGGCCACGAACTGATGAGATATGTCACCCGAGACGATGGCCA	180
Db	308	GTGGCTGTCTCAACCAACCAAGGCCACGAACTGATGAGATATGTCACCCGAGACGATGGCCA	367
QY	181	ACCTGGCCGGTAAAACTGGTGTGGCTGACACCGCGAGTGGTGAAGCTAATCGCCATATCTC	240
Db	368	ACCTGGCCGGTAAAACTGGTGTGGCTGACACCGCGAGTGGTGAAGCTAATCGCCATATCTC	427
QY	241	GGACATTACCAACGGCCAGCGTCTGGTAGCCGACGAGATGCTGGCGAAAAAGGGCTGCCG	300
Db	428	GGACATTACCAACGGCCAGCGTCTGGTAGCCGACGAGATGCTGGCGAAAAAGGGCTGCCG	487
QY	301	CAGTGTGCTCAGGTGCGCATTAATGCTGATGTCGACGACGCTACCCGTAACAGCGCA	360
Db	488	CAGTGTGCTCAGGTGCGCATTAATGCTGATGTCGACGACGCTACCCGTAACAGCGCA	547
QY	361	GCGTTGCGCGCGGGCTGGCACTGACGTGCATTAATACCGTAATACCCAGGACGATACG	420
Db	548	GCGTTGCGCGCGGGCTGGCACTGACGTGCATTAATACCGTAATACCCAGGACGATACG	607
QY	421	TCGACGTCCCGATCCGTTATTTAATCCCTAAAAAACCTGGCGTTTGGCAACTGGATTAACGCG	480
Db	608	TCGACGTCCCGATCCGTTATTTAATCCCTAAAAAACCTGGCGTTTGGCAACTGGATTAACGCG	667
QY	481	AACGTGACTACGCGCATCTCTACGACGGCGACGAGAGGGTAATTGTCTGACTTTACGGGCAT	540
Db	668	AACGTGACTACGCGCATCTCTACGACGGCGACGAGAGGGTAATTGTCTGACTTTACGGGCAT	727
QY	541	CGGCAAAACGGGGGTTTCGGGAAGTGGAAAGGGGGCTTAATTTTCCGCATCAAACTGTGCG	600
Db	728	CGGCAAAACGGGGGTTTCGGGAAGTGGAAAGGGGGCTTAATTTTCCGCATCAAACTGTGCG	787
QY	601	CTTAAACGTGAGAAAACAGAGCAAAAGCTGTTTCATTACGACGCAATTACATCGGAATCTC	660
Db	788	CTTAAACGTGAGAAAACAGAGCAAAAGCTGTTTCATTACGACGCAATTACATCGGAATCTC	847
QY	661	AAGGTAGACCGCCGACATATCTCATTTAACCGCGTGGGTAAAGCTTGACATATGTGTACG	720
Db	848	AAGGTAGACCGCCGACATATCTCATTTAACCGCGTGGGTAAAGCTTGACATATGTGTACG	907
QY	721	GAGATATTTTCCTCTCAACAGCAACAGGAAATGCCGAGACCGGGGTGGGGAAGATATCAC	780
Db	908	GAGATATTTTCCTCTCAACAGCAACAGGAAATGCCGAGACCGGGGTGGGGAAGATATCAC	967
QY	781	GATTCACACACAGTGGACACCTGTGCTAATGTTTGCATTAACGGCGCAATTTTATTTGGCTACAA	840
Db	968	GATTCACACACAGTGGACACCTGTGCTAATGTTTGCATTAACGGCGCAATTTTATTTGGCTACAA	1027
QY	841	CGCACGCCAGAGTGGCCCCGACGCCGCCACCCCGTTATTTGATTTGATCATGTGGCACGCG	900
Db	1028	CGCACGCCAGAGTGGCCCCGACGCCGCCACCCCGTTATTTGATTTGATCATGTGGCACGCG	1087
QY	901	TTGACGCCCCCATCCACCGGCAAAAACAGCGCGATGGTGTGACATTTACCACTCCAGTACTG	960
Db	1088	TTGACGCCCCCATCCACCGGCAAAAACAGCGCGATGGTGTGACATTTACCACTCCAGTACTG	1147
QY	961	TTTATATGCGGACACAGTACTAATCTGGCAAAATCTCGCGCGCGCACTGGACCTCAACTGCG	1020
Db	1148	TTTATATGCGGACACAGTACTAATCTGGCAAAATCTCGCGCGCGCACTGGACCTCAACTGCG	1207
QY	1021	ACGCTTCCCGGTCAACGCGATPAACAGCCAGTGAATTAAGTTTGCTGCTTCACAACCTTTAAG	1080
Db	1208	ACGCTTCCCGGTCAACGCGATPAACAGCCAGTGAATTAAGTTTGCTGCTTCACAACCTTTAAG	1267
QY	1081	CGTGGCGTAACGAGTAACAGCCAGTGAATTAAGTTTGCTGCTTCACAACCTTTAAG	1146
Db	1268	CGTGGCGTAACGAGTAACAGCCAGTGAATTAAGTTTGCTGCTTCACAACCTTTAAG	1327
QY	1141	CAGATGCGTGAATTAACAGCCGCTGTGCATTAAATAGCGCCCGGAGAGAGTAAACTGAC	1200
Db	1328	CAGATGCGTGAATTAACAGCCGCTGTGCATTAAATAGCGCCCGGAGAGAGTAAACTGAC	1387
QY	1201	CTGGAGGATGTGAAGACGAAATGCGCAGGGCATGTGTCGTTGGCAGGTTTAAACGCA	1260

DB 1388 CTGGGAGGATGTGAAGAGCAATGCGGAGCATGTCTTGGCAGATTTCAGCAA 1447
OY 1261 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAGAT 1300
DB 1448 ATCGTAATGAAGACGATACCGGCGTGCAGTTTGAT 1487

RESULT 14
ECOCAPHC 1901 bp DNA linear BCT 04-APR-2002
LOCUS Escherichia coli periplasmic phosphoanhydride phosphohydrolase
DEFINITION (appa) gene, complete cds.
ACCESSION L03372.1 GI:145180
VERSION L03372.1
KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
SOURCE Escherichia coli.
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE 1 (sites)
AUTHORS Dassa, J., Marck, C. and Boquet, P.L.
TITLE The complete nucleotide sequence of the Escherichia coli gene appa
reveals significant homology between pH 2.5 acid phosphatase and
glucose-1-phosphatase
J. Bacteriol. 172 (9), 5497-5500 (1990)
JOURNAL 90368616
MEDLINE 2168385
PUBMED
REFERENCE 2 (bases 1 to 1901)
AUTHORS Ostalan, K., Harms, E.H., Stevils, P.E., Kuciel, R., Zhou, M.M. and Van
Etten, R.L.
TITLE Overexpression, site-directed mutagenesis, and mechanism of
Escherichia coli acid phosphatase
J. Biol. Chem. 267 (32), 22830-22836 (1992)
JOURNAL 93054596
MEDLINE 1429631
PUBMED

FEATURES
source
1. location/Qualifiers
1.1901
/organism="Escherichia coli"
/db_xref="taxon:562"
188.1486
/gene="appa"
188.1486
/gene="appa"
/standard_name="acid phosphatase"
/citation-[1]
/codon_start=1
/transl_table=11
/product="periplasmic phosphoanhydride phosphohydrolase"
/protein_id="AA00004.1"
/db_xref="GI:145181"
/translation="MKAILIPLSLILPLPOSAPFAOSEPELLESVIVSHGVAAP
TKATQIMODVTPDAMPVVKLGMPLPRGGLIAYIGHYORORLVAADGLAKGPOS
GOVAIIADVDERTRKTEGAPFAAGLAPOCAITVHOADISDPDLFNPILKTGVCOLDA
NVTDAIILRAGSGINDPYGHOTARRELELYNFPQSNICLKRQEDSCSLTQALPS
ELKVSADVSLGAVSLASMLTEIFLLQDQMPPEPGRITDSHOWMTLSLHNAOF
YLELMTPLPGQPDNTPGGEIVFERMRRLSDSONIOWSLVFOFLQOMRDKPLSLNT
PPGEVKLTIAGEEERNAGCMGSLAGFTQIVNARIAPASL"
299.301
variation
/gene="appa"
/citation-[1]
/phenotype="R20A mutant"
BASE COUNT 474 a 499 c 499 g 428 t 1 others
ORIGIN

Query Match 97.1%; Score 1284.6; DB 1; Length 1901;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 1291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;

OY 1 ATGAAGGAGACTTAATCCATTTTATCTCTTGATTCGTTAAACCCGCAATCGA 60
DB 188 ATGAAGGAGACTTAATCCATTTTATCTCTTGATTCGTTAAACCCGCAATCGA 247

OY 61 TTCGCTCAGAGTAGAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTACGTGTCATGCT 120
DB 248 TTCGCTCAGAGTAGAGCCGAGCTGAAGCTGAAAGTGTGTGATTTGTACGTGTCATGCT 307
OY 121 GTGCGTCTCCCAACCAAGGCCACGACATGATGACAGATGTACACCCAGACGATGGCA 180
DB 308 GTGGCTGCTCAACCAAGGCCACGACATGATGACAGATGTACACCCAGACGATGGCA 367
OY 181 ACCTGGCCGGTAAACTGGGTGGCTGACACCCGCGNGGTGGTACCTAATCCCTATCTC 240
DB 368 ACCTGGCCGGTAAACTGGGTGGCTGACACCCGCGNGGTGGTACCTAATCCCTATCTC 427
OY 241 GGACATTTACCAAGCCAGCGTGTGTAGCCGACGATTTGTGGCGAAAAAGGGCTGCCG 300
DB 428 GGACATTTACCAAGCCAGCGTGTGTAGCCGACGATTTGTGGCGAAAAAGGGCTGCCG 487
OY 301 CAGTCTGGTCAAGTGGCGGATTTATGTGTATGTGACAGACAGCTACCCGTTAAACAGCGAA 360
DB 488 CAGTCTGGTCAAGTGGCGGATTTATGTGTATGTGACAGACAGCTACCCGTTAAACAGCGAA 547
OY 361 GCCTTCGCGCGCGGCTGGCACCCTGACCTGCAATTAACCGTACATACCCGACAGATAG 420
DB 548 GCCTTCGCGCGCGGCTGGCACCCTGACCTGCAATTAACCGTACATACCCGACAGATAG 607
OY 421 TCCAGTCCCGATCCGTTATTTAATCCCTAATAAAGTGGGTTGGCCAACTGGATTAACGG 480
DB 608 TCCAGTCCCGATCCGTTATTTAATCCCTAATAAAGTGGGTTGGCCAACTGGATTAACGG 667
OY 481 AACGTGACAGACCGCATCTCTACAGACGAGGAGGATCAATTTGCTGATTTACGGGCA 540
DB 668 AACGTGACAGACCGCATCTCTACAGACGAGGAGGATCAATTTGCTGATTTACGGGCA 727
OY 541 CGGCAAAAGCGGCTTTCGCAACTGGAACGGGCTTAATTTCCCAATCAACTGTGTC 600
DB 728 CGGCAAAAGCGGCTTTCGCAACTGGAACGGGCTTAATTTCCCAATCAACTGTGTC 787
OY 601 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGATTTACATGCAATC 660
DB 788 CTTAAACGTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGATTTACATGCAATC 847
OY 661 AAGGTGAGCGCGACAAATGTCATTAACGGTGGGCTTAAGCTGCAATGCTGAGAG 720
DB 848 AAGGTGAGCGCGACAAATGTCATTAACGGTGGGCTTAAGCTGCAATGCTGAGAG 907
OY 721 GAGATATTTCTCCTGCAACACAGAGGAATCCGAGCCGGGTGGGGAAGATACAC 780
DB 908 GAGATATTTCTCCTGCAACACAGAGGAATCCGAGCCGGGTGGGGAAGATACAC 967
OY 781 GATTCACACCACTGGAAACCTTGTCTAAGTTGCAATACGGCAATTTATTTGCTACA 840
DB 968 GATTCACACCACTGGAAACCTTGTCTAAGTTGCAATACGGCAATTTATTTGCTACA 1027
OY 841 CGCAGCCGAGAGTGGCCCGGACCGCGCCGACCCGCTTAAGTTGATTCATGTCAGCG 900
DB 1028 CGCAGCCGAGAGTGGCCCGGACCGCGCCGACCCGCTTAAGTTGATTCATGTCAGCG 1087
OY 901 TTGACGCCCATCCACCGCAAAAGAGCGGTATGTTGACATTAACCCACTCACTACTG 960
DB 1088 TTGACGCCCATCCACCGCAAAAGAGCGGTATGTTGACATTAACCCACTCACTACTG 1147
OY 961 TTTATTCGCGGACAGATTAATCTGGCAAAATCTGGCGGCGCACTGGAGCTCAATCG 1020
DB 1148 TTTATTCGCGGACAGATTAATCTGGCAAAATCTGGCGGCGCACTGGAGCTCAATCG 1207
OY 1021 ACGCTTCCCGGTCACCGGATTAACACGCCCGCAGGTGTGATGCTGTTGAAGCGTGG 1080
DB 1208 ACGCTTCCCGGTCACCGGATTAACACGCCCGCAGGTGTGATGCTGTTGAAGCGTGG 1267
OY 1081 CGTGGCTAAGGAGTAACAGCAGTGTGATTCAGTTTGTGCTGCTTCCAGACTTTAAG 1140
DB 1268 CGTGGCTAAGGAGTAACAGCAGTGTGATTCAGTTTGTGCTGCTTCCAGACTTTAAG 1327

Oy 1141 CAGATGGCTGATTAAGCCGCTGTCAATTAATAGCCGCCGAGAGGTAAATGACC 1200
 |||||||
 Db 1328 CAGATGGCTGATTAAGCCGCTGTCAATTAATAGCCGCCGAGAGGTAAATGACC 1387
 |||||||
 Oy 1201 CTGGCAGGATGTGAAGAGCCAAATGCGAGGATGTTCGTTGGCAGGTTTACCAA 1260
 |||||||
 Db 1388 CTGGCAGGATGTGAAGAGCCAAATGCGAGGATGTTCGTTGGCAGGTTTACCAA 1447
 |||||||
 Oy 1261 ATGCTGAATGAGCAGCATACCGGCGTGCAGTTTGAGAT 1300
 |||||||
 Db 1448 ATGCTGAATGAGCAGCATACCGGCGTGCAGTTTGAT 1487
 |||||||

RESULT 15
 ECOMAPHD 1901 bp DNA linear BCT 04-APR-2002
 LOCUS
 DEFINITION Escherichia coli periplasmic phosphoanhydride phosphohydrolase
 (appa) gene, complete cds.
 ACCESSION L03373.1 GI:145182
 VERSION L03373
 KEYWORDS acid phosphatase; periplasmic phosphoanhydride phosphohydrolase.
 SOURCE Escherichia coli.
 ORGANISM Escherichia.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE 1 (sites)
 AUTHORS Dassa, J., Marck, C. and Boquet, P. L.
 TITLE The complete nucleotide sequence of the Escherichia coli gene appa reveals a significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase
 JOURNAL J. Bacteriol. 172 (9), 5497-5500 (1990)
 MEDLINE 90368616
 PUBMED 2168385
 2 (bases 1 to 1901)
 Ostalan, R., Harms, E. H., Stevis, P. E., Kuciel, R., Zhou, M. M. and Van Etten, R. L.
 TITLE Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase
 JOURNAL J. Biol. Chem. 267 (32), 22830-22836 (1992)
 MEDLINE 93054596
 PUBMED 1429631

FEATURES
 source location/Qualifiers
 1..1901 /organism="Escherichia coli"
 /db_xref="taxon:562"
 188..1486 /gene="appa"
 /gene="appa"
 /standard_name="acid phosphatase"
 /citation-[1]
 /codon_start=1
 /transl_table=1
 /product="periplasmic phosphoanhydride phosphohydrolase"
 /protein_id="AA00005.1"
 /db_xref="GI:145182"
 /translation="MKAILIPELSLIPITPQSAFQSEPEIKIESVIVSRHGVAP
 TKATQIMQDTPDMPTWPKYKIGLTPRGELIATLGHQRLVAGLAKKCPQS
 GOVAIADVBERKTEAFAGIAPCAITVHOADTSSPDLFNLKGTGVCGLDA
 NVTDAISRAGSIADETGHROTAFRELERYLNPOSNLCIKREKODESCLQALPS
 ELVYASDNVSLTGAVSLASMLTEIFLQQAQMPGEMGRITDSHONNTLSLNAPF
 YLORTEPEVARSRAITPLDLIKTALTPHPQOKAYGVTLPSVFIAGHDPNLANLG
 ALRNMTLPQDPNTPRGELVPERMRSLDSNOMIOVSLVFOIQQMRDKTPISLNTG
 PRGVKTLTLAGCERNAGMCISLAGFTOIYENARITPACSL"
 299..301 /gene="appa"
 /citation-[1]
 /phenotype="R16a mutant"
 variation /base_count="499 a 499 c 499 g 428 t 1 others"

BASE COUNT 474 a 499 c 499 g 428 t 1 others
 ORIGIN
 Query Match 97.1%; Score 1284.6; DB 1; Length 1901;
 Best Local Similarity 99.3%; Pred. No. 0;

Matches 1291; Conservative 0; Mismatches 9; Indels 0; Gaps 0;
 Oy 1 ATGAAGGATCTTAATCCATTTTATCTTGTATTCCTTAACCCGCAATTCGA 60
 |||||||
 Db 188 ATGAAGGATCTTAATCCATTTTATCTTGTATTCCTTAACCCGCAATTCGA 247
 |||||||
 Oy 61 TTGGCTAGAGTGGCCGGAGCTAGAGTGAAGTGGTGGATGTGACGTCGATAGT 120
 |||||||
 Db 248 TTGGCTAGAGTGGCCGGAGCTAGAGTGAAGTGGTGGATGTGACGTCGATAGT 307
 |||||||
 Oy 121 GTGGGTGTCCCAACCAAGGCGCACACTGATGAGATGTCACCCAGAGATGGCCA 180
 |||||||
 Db 308 GTGGGTGTCCCAACCAAGGCGCACACTGATGAGATGTCACCCAGAGATGGCCA 367
 |||||||
 Oy 181 ACTTGGCCGCTAAACTGGTGGCTGACACCGCGNGGTGTGAGTAACTGCTATCTC 240
 |||||||
 Db 368 ACTTGGCCGCTAAACTGGTGGCTGACACCGCGNGGTGTGAGTAACTGCTATCTC 427
 |||||||
 Oy 241 GGACATTAACCAAGCGCAGCGCTGTGTAAGCCGAGGATTCGCGGAAAGGCGTGGCCG 300
 |||||||
 Db 428 GGACATTAACCAAGCGCAGCGCTGTGTAAGCCGAGGATTCGCGGAAAGGCGTGGCCG 487
 |||||||
 Oy 301 CAGTGTGTGAGTGTGCGGATTAATGCTGATGTGACGAGCGTACCCGTTAAACAGCGCAA 360
 |||||||
 Db 488 CAGTGTGTGAGTGTGCGGATTAATGCTGATGTGACGAGCGTACCCGTTAAACAGCGCAA 547
 |||||||
 Oy 361 GCTTTCGCGCGCGGGCTGGACCTGATGTGCAATTAACCGTACATCCAGGCAATACG 420
 |||||||
 Db 548 GCTTTCGCGCGCGGGCTGGACCTGATGTGCAATTAACCGTACATCCAGGCAATACG 607
 |||||||
 Oy 421 TCAGTCCCGATTCGCTTATTAATCCTTAATAAAGCGCTTCCCAACGTGATACAGCG 480
 |||||||
 Db 608 TCAGTCCCGATTCGCTTATTAATCCTTAATAAAGCGCTTCCCAACGTGATACAGCG 667
 |||||||
 Oy 481 AACGTGACTGACGCGATCTTCAGCAGGCGCAGAGGCTCAATGCTGACTTACCGGGCAT 540
 |||||||
 Db 668 AACGTGACTGACGCGATCTTCAGCAGGCGCAGAGGCTCAATGCTGACTTACCGGGCAT 727
 |||||||
 Oy 541 CGGCAACGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 600
 |||||||
 Db 728 CGGCAACGGCGTTTCGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAACTTGTGC 787
 |||||||
 Oy 601 CTTAAAGGTGAACAGAGGAGCAAGAGCTTCAATTAACGCGGATTAACCATTCGGAATC 660
 |||||||
 Db 788 CTTAAAGGTGAACAGAGGAGCAAGAGCTTCAATTAACGCGGATTAACCATTCGGAATC 847
 |||||||
 Oy 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 720
 |||||||
 Db 848 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAGCTCGCATCAATGCTGACG 907
 |||||||
 Oy 721 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAACCGGGGTGGGGAAGATCAC 780
 |||||||
 Db 908 GAGATATTTCTCTGCAACAAGCACAGGGAATGCCGGAACCGGGGTGGGGAAGATCAC 967
 |||||||
 Oy 781 GATTACACACAGTGGAAACACCTTCTAAGTTTGATTAACGGCAATTTTATTTCTCAAA 840
 |||||||
 Db 968 GATTACACACAGTGGAAACACCTTCTAAGTTTGATTAACGGCAATTTTATTTCTCAAA 1027
 |||||||
 Oy 841 CGCAGCGCAGAGTGTGCCCGAGCGCGCACCCCGTTATTTGATTTGATCAGTCAAGC 900
 |||||||
 Db 1028 CGCAGCGCAGAGTGTGCCCGAGCGCGCACCCCGTTATTTGATTTGATCAGTCAAGC 1087
 |||||||
 Oy 901 TTGACGCGCCATCCACGCGCAAAACAGGCGTATGCTGACATTAACCATCTTCAGTAC 960
 |||||||
 Db 1088 TTGACGCGCCATCCACGCGCAAAACAGGCGTATGCTGACATTAACCATCTTCAGTAC 1147
 |||||||
 Oy 961 TTTATTTGCGGACAGATTAATCTGGCAATTCGCGGCGGACTGAGACTCAACTGG 1020
 |||||||
 Db 1148 TTTATTTGCGGACAGATTAATCTGGCAATTCGCGGCGGACTGAGACTCAACTGG 1207
 |||||||
 Oy 1021 ACGCTTCCCGGTGACCGGATTAACACCGCGCAGGTGATGTAATGATTTGAACGCTGG 1080
 |||||||
 Db 1208 ACGCTTCCCGGTGACCGGATTAACACCGCGCAGGTGATGTAATGATTTGAACGCTGG 1267
 |||||||

QY 1081 CGTCGGCTAAGCATTAACAGCCAGTGCATTTCAGGTTTCGCTGCTCTTCCAGACTTTCAG 1140
 |||||
 Db 1268 CGTCGGCTAAGCATTAACAGCCAGTGCATTTCAGGTTTCGCTGCTCTTCCAGACTTTCAG 1327
 |||||
 QY 1141 CAGATGCCGTGATTAACAGCCGCTGTCATTAAATACGCCGCCCGAGAGGTGAACACTGACC 1200
 |||||
 Db 1328 CAGATGCCGTGATTAACAGCCGCTGTCATTAAATACGCCGCCCGAGAGGTGAACACTGACC 1387
 |||||
 QY 1201 CTGGCAGATGTGAAGAGCGAATGGCAGAGGCGATGTTCGTTGGCAGGTTTACGCAA 1260
 |||||
 Db 1388 CTGGCAGATGTGAAGAGCGAATGGCAGAGGCGATGTTCGTTGGCAGGTTTACGCAA 1447
 |||||
 QY 1261 ATCGTGAATGAAGACAGCATACCGGCGTGCAGTTGAGAT 1300
 |||||
 Db 1448 ATCGTGAATGAAGACAGCATACCGGCGTGCAGTTGTAAT 1487
 |||||

Search completed: June 12, 2003, 11:33:25
 Job time : 3573 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 12, 2003, 09:19:38 ; Search time 2118 Seconds
(Without alignments)
10116.444 Million cell updates/sec

Title: US-09-777-566A-1
Perfect score: 1323
Sequence: 1 atgaagcgccttaaccc.....atcacatcacatcactaa 1323

Scoring table: IDENTITY_NUC
Gapop.10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST: *
1: em_estbda:*
2: em_esthum:*
3: em_estln:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hnc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hnc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	614.4	46.4	616	BJ074127	BJ074127 BJ074127
2	386.2	29.2	393	BE520240	BE520240 M1B12STM
3	383	28.9	383	AM036132	AM036132 EST274508
4	383	28.9	383	AM036134	AM036134 EST274510
5	304.6	23.0	354	AA545747	AA545747 HBMSF1B4-
6	216.8	16.4	320	BE436403	BE436403 EST407481

Result No.	Score	Query Match	Length	DB ID	Description
7	208	15.7	208	BM409016	BM409016 EST583343
8	208	15.7	208	BM412806	BM412806 EST587133
9	198.8	15.0	546	BJ040796	BJ040796 BJ040796
10	168.4	12.8	179	BE461872	BE461872 EST413291
11	160.8	12.2	508	BJ030203	BJ030203 BJ030203
12	157.6	11.9	657	BE457132	BE457132 NF10C05P
13	143.2	10.8	392	BE520241	BE520241 M1B12XTM
14	43.8	3.3	354	14	BE520241
15	41.2	3.1	733	13	BE520241
16	39.8	3.0	431	9	BE520241
17	39.8	3.0	602	10	BE520241
18	39.8	3.0	607	9	BE520241
19	39.2	3.0	553	10	BE520241
20	38.6	2.9	398	9	BE520241
21	38.6	2.9	669	17	BE520241
22	38.2	2.9	379	10	BE520241
23	38.2	2.9	459	9	BE520241
24	38.2	2.9	693	14	BE520241
25	38.2	2.9	693	14	BE520241
26	38.2	2.9	693	14	BE520241
27	38.2	2.9	694	10	BE520241
28	38.2	2.9	756	14	BE520241
29	38.2	2.9	900	14	BE520241
30	37.8	2.9	942	13	BE520241
31	37.6	2.8	274	14	BE520241
32	37.6	2.8	553	10	BE520241
33	36.8	2.8	618	17	BE520241
34	36.4	2.8	875	12	BE520241
35	36.2	2.7	674	17	BE520241
36	36.2	2.7	764	12	BE520241
37	36	2.7	638	13	BE520241
38	36	2.7	790	17	BE520241
39	35.8	2.7	957	12	BE520241
40	35.6	2.7	396	9	BE520241
41	35.4	2.7	238	10	BE520241
42	35.4	2.7	295	9	BE520241
43	35.4	2.7	389	9	BE520241
44	35.4	2.7	475	9	BE520241
45	35.4	2.7	483	10	BE520241

ALIGNMENTS

RESULT 1
BJ074127/c 616 bp mRNA linear EST 11-DEC-2001
LOCUS BJ074127 NIBB Mochi1 normalized Xenopus tailbud library Xenopus
DEFINITION laevis cDNA clone X1090j06 5', mRNA sequence.
ACCESSION BJ074127
VERSION BJ074127.1 GI:17504316
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
REFERENCE 1 (bases 1 to 616)
Xenopus laevis.
Xenopodinae; Xenopus.
Kittayama, A., Terasaka, C., Mochi, M., Ueno, N., Shin, I., T. and Kohara
Y.
Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadatsu Shin-I
Center for Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshtn@genes.nig.ac.jp.
Location/Qualifiers
1..616
/organism="Xenopus laevis"
/db_xref="taxon:8355"

/clone="XL090j06"
 /clone_lib="NIBB Mochli normalized Xenopus tailbud
 library"
 /issue_type="whole embryo"
 /dev_stage="stage 25"
 /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAS
 were oligo-dt primed and directionally cloned. Staging
 according to Nieuwkoop and Faber. Library is subcloned
 and was constructed by N. Garrett and A.M. Zorn,
 (Wellcome/CRC Institute)."

BASE COUNT 128 a 167 c 167 g 154 t
 ORIGIN

Query Match 46.4%; Score 614.4; DB 13; Length 616;
 Best Local Similarity 99.8%; Pred. No. 5.1e-190;
 Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 264 GGTACCCGACGATCTGCGCAAAAAGGCTCCCGCACTGCGTACAGTCCGATAT 323
 DB 616 GGTACCCGACGATCTGCGCAAAAAGGCTCCCGCACTGCGTACAGTCCGATAT 557
 QY 324 TGCTGATGTCAGACGATACCGCTAAACAGGCGAAGCTTCGCGCGCTGGCACC 383
 DB 556 TGCTGATGTCAGACGATACCGCTAAACAGGCGAAGCTTCGCGCGCTGGCACC 497
 QY 384 TGAAGTGCATTAACGCTACATACCGACGATAGTCCAGTCCGATCCGTTATTAA 443
 DB 496 TGAAGTGCATTAACGCTACATACCGACGATAGTCCAGTCCGATCCGTTATTAA 437
 QY 444 TCCTCTAAACAGTGGGCTTGGCAACTGATGATACGCGACGATGATGATGATGAT 503
 DB 436 TCCTCTAAACAGTGGGCTTGGCAACTGATGATGATGATGATGATGATGATGAT 377
 QY 504 CAGGCGACGAGGCTCAATGCTGATTTACCGGCGATCGGCAACGGCGTTTGGCACT 563
 DB 376 CAGGCGACGAGGCTCAATGCTGATTTACCGGCGATCGGCAACGGCGTTTGGCACT 317
 QY 564 GGAACGGGCTTAAATTTTCCGCAATCAACTTGTGCTTAAAGTGAAGAAAGCA 623
 DB 316 GGAACGGGCTTAAATTTTCCGCAATCAACTTGTGCTTAAAGTGAAGAAAGCA 257
 QY 624 AAGCTGTCATTAACGAGGATTAACGATGATGATGATGATGATGATGATGATGAT 683
 DB 256 AAGCTGTCATTAACGAGGATTAACGATGATGATGATGATGATGATGATGATGAT 197
 QY 684 ATTAACCGGTGCGGTAAGCTCGCATCAATGCTGACGAGATATTTCTCTGCAACAAC 743
 DB 196 ATTAACCGGTGCGGTAAGCTCGCATCAATGCTGACGAGATATTTCTCTGCAACAAC 137
 QY 744 ACAGGGAATGCGGAGCGGGGTGGGAAGATCAACGATTCACACCAAGTGAACACTT 803
 DB 136 ACTGGGAATGCGGAGCGGGGTGGGAAGATCAACGATTCACACCAAGTGAACACTT 77
 QY 804 GCTAAGTTGTCATTAACGCGCAATTTTATTTGCTCAACAGCAGCAGAGTGGCCGAG 863
 DB 76 GCTAAGTTGTCATTAACGCGCAATTTTATTTGCTCAACAGCAGCAGAGTGGCCGAG 17
 QY 864 CCGCGCACCCCGTTA 879
 DB 16 CCGCGCACCCCGTTA 1

RESULT 2
 BE520240/c 393 bp mRNA linear EST 19-MAR-2001
 LOCUS M1B12STM Arabidopsis developing seed Arabidopsis thaliana cDNA
 DEFINITION clone M1B12 5', mRNA sequence.

ACCESSION BE520240
 VERSION BE520240.1 GI:9778242
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 COMMENT
 Spectrophobia; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 1 (bases 1 to 393)
 White, J.A., Todd, J., Newman, T., Focks, N., Girke, T., Martinez de
 Illarduya, O., Jaworski, J.G., Ohlrogge, J., and Bening, C.
 A new set of Arabidopsis expressed sequence tags from developing
 seeds. The metabolic pathway from carbohydrates to seed oil
 plant physiol. 124 (4), 1582-1594 (2000)
 20567808
 Contact: Bening, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu

FEATURES
 source
 Location/Qualifiers
 1..393
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone_lib="M1B12"
 /issue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site_1: EcoRI; Site_2: XhoI"

BASE COUNT 86 a 100 c 103 g 102 t 2 others
 ORIGIN

Query Match 29.2%; Score 386.2; DB 10; Length 393;
 Best Local Similarity 98.7%; Pred. No. 3.5e-115;
 Matches 388; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 319 ATTAATTCCTGATGTCGACGAGCTGATCCGTTAAACAGGCGAAGCTTCGCGCGGGCTG 378
 DB 393 ATTAATTCCTGATGTCGACGAGCTGATCCGTTAAACAGGCGAAGCTTCGCGCGGGCTG 334
 QY 379 GCACCTGACGTGCAATTAACGCTACATACCGAGCAGATACGTCAGTCCGATCCGTTA 438
 DB 333 GCACCTGACGTGCAATTAACGCTACATACCGAGCAGATACGTCAGTCCGATCCGTTA 274
 QY 439 TTATATCCTCTAAACACTGCGGCTTGGCAACTGATTAACGCGAAGCTGACGCGATC 498
 DB 273 TTATATCCTCTAAACACTGCGGCTTGGCAACTGATTAACGCGAAGCTGACGCGATC 214
 QY 499 CTCAGCAGGCGAGAGGCTCAATGTCGACTTACCGGCGCATCGGCAACGGCGTTGCG 558
 DB 213 CTCAGCAGGCGAGAGGCTCAATGTCGACTTACCGGCGCATCGGCAACGGCGTTGCG 154
 QY 559 GAACGTGAAGGCGCTTAATTTTCCGCAATCAACTGTCCTTAAACGTGAGAAACAG 618
 DB 153 GAACGTGAAGGCGCTTAATTTTCCGCAATCAACTGTCCTTAAACGTGAGAAACAG 94
 QY 619 GACGAACCTGTCATTAACGAGGATTAACGATGATGATGATGATGATGATGATGAT 678
 DB 93 GACGAACCTGTCATTAACGAGGATTAACGATGATGATGATGATGATGATGATGAT 34
 QY 679 GTCTCATTAACCGGTGCGGTAAGCTTCGCATCA 711
 DB 33 GTCTCATTAACCGGTGCGGTAAGCTTCGCATCA 1

RESULT 3
 AW036132 383 bp mRNA linear EST 18-MAY-2001
 LOCUS AW036132
 DEFINITION EST74508 tomato seed, TAMU lycopersicon esculentum cDNA clone

CLEIE23 similar to periplasmic phosphoanhydride phosphohydrolase precursor, putative, mRNA sequence.

ACCESSION
AM036132
KEYWORDS
AM036132.1 GI:5894811
SOURCE
EST.
ORGANISM
tomato.

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
1 (bases 1 to 383)

Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato seed tissue
Unpublished (1999)

TITLE

JOURNAL

COMMENT

CONTACT

LOCATION

EMAIL

URL

FEATURES

SOURCE

1. 383

/organism="Lycopersicon esculentum"

/cultivar="TA96"

/db_xref="taxon:4081"

/clone="CLEIE23"

/clone_1db="tomato seed, TAMU"

/tissue_type="seeds"

/dev_stage="quiescent seed"

/lab_host="X11-Blue MR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; CLEIE - Tomato Seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."

BASE COUNT

ORIGIN

102 a 101 c 97 g 83 t

Query Match

Best Local Similarity

Matches

383; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

28.9%; Score 383; DB 10; Length 383;

Pred. No. 3.9e-114;

1 TAAACAGGCGAAGCCTTCGCCGCCGCGGCTGGCACCTGCTCAATTAACCGTACATAC

60

408 CCAAGGAGATACGTCGATCCGCGGCTGATTTAATTCCTTAAACTGGCGTTGGCA

467

61 CCAAGGAGATACGTCGATCCGCGGCTGATTTAATTCCTTAAACTGGCGTTGGCA

120

468 ACTGATTAACGCGAAGCTGATGACGCGATCTTCAGCAGGAGGAGGTCATTTGCTGA

527

121 ACTGATTAACGCGAAGCTGATGACGCGATCTTCAGCAGGAGGAGGTCATTTGCTGA

180

528 CTTTACCGGCGATCGGCAAAAGCGGCTTCGCGAATCGGAACGGGTGCTTAATTTCCGCA

587

181 CTTTACCGGCGATCGGCAAAAGCGGCTTCGCGAATCGGAACGGGTGCTTAATTTCCGCA

240

588 ATCAAACTGTGCTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATC

647

241 ATCAAACTGTGCTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATC

300

648 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGGTGAACCTCGC

707

301 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGGTGAACCTCGC

360

708 ATCAATGCTGACGAGATATTTTC

730

361 ATCAATGCTGACGAGATATTTTC

383

RESULT 4
AM036134/c
LOCUS
DEFINITION
EST274510 tomato seed, TAMU Lycopersicon esculentum cDNA clone

CLEIE23 similar to periplasmic phosphoanhydride phosphohydrolase precursor, putative, mRNA sequence.

ACCESSION
AM036134
KEYWORDS
AM036134.1 GI:5894813
SOURCE
EST.
ORGANISM
tomato.

Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.

REFERENCE
1 (bases 1 to 383)

Alcala, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E., Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman, C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley, S.D. and Giovannoni, J.

Generation of ESTs from tomato seed tissue
Unpublished (1999)

TITLE

JOURNAL

COMMENT

CONTACT

LOCATION

EMAIL

URL

FEATURES

SOURCE

1. 383

/organism="Lycopersicon esculentum"

/cultivar="TA96"

/db_xref="taxon:4081"

/clone="CLEIE23"

/clone_1db="tomato seed, TAMU"

/tissue_type="seeds"

/dev_stage="quiescent seed"

/lab_host="X11-Blue MR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; CLEIE - Tomato Seed EST library. Directionally cloned cDNAs inserted into pBluescript SK(-) at 5' end with EcoRI and 3' end with XhoI site."

BASE COUNT

ORIGIN

83 a 97 c 101 g 102 t

Query Match

Best Local Similarity

Matches

383; Conservative

0; Mismatches

0; Indels

0; Gaps

0;

28.9%; Score 383; DB 10; Length 383;

Pred. No. 3.9e-114;

348 TAAACAGGCGAAGCCTTCGCCGCCGCGGCTGGCACCTGCTCAATTAACCGTACATAC

407

383 TAAACAGGCGAAGCCTTCGCCGCCGCGGCTGGCACCTGCTCAATTAACCGTACATAC

324

408 CCAAGGAGATACGTCGATCCGCGGCTGATTTAATTCCTTAAACTGGCGTTGGCA

467

323 CCAAGGAGATACGTCGATCCGCGGCTGATTTAATTCCTTAAACTGGCGTTGGCA

264

468 ACTGATTAACGCGAAGCTGATGACGCGATCTTCAGCAGGAGGAGGTCATTTGCTGA

527

263 ACTGATTAACGCGAAGCTGATGACGCGATCTTCAGCAGGAGGAGGTCATTTGCTGA

204

528 CTTTACCGGCGATCGGCAAAAGCGGCTTCGCGAATCGGAACGGGTGCTTAATTTCCGCA

587

203 CTTTACCGGCGATCGGCAAAAGCGGCTTCGCGAATCGGAACGGGTGCTTAATTTCCGCA

144

588 ATCAAACTGTGCTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATC

647

143 ATCAAACTGTGCTTAAAGCTGAGAAACAGAGAAAGCTGTCATTAACGCGGATC

84

648 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGGTGAACCTCGC

707

83 ACCATCGGAACCTCAAGGTGAGCGCGCAATGTCTCTTAACGCGGTGAACCTCGC

24

OY 708 ATCAATGCTGACGAGATATTC 730
 DB 23 ATCAATGCTGACGAGATATTC 1

RESULT 5 354 bp mRNA 11near EST 12-MAY-1999
 AA545747/c
 LOCUS HBMSFLB4-REV Human Bone Marrow Stromal Fibroblast Homo sapiens CDNA
 DEFINITION clone HBMSFLB4 5', mRNA sequence.
 ACCESSION AA545747
 VERSION AA545747
 KEYWORDS GI:2307026
 SOURCE EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 354)
 Jia, L., Robey, P., Young, M., Schuler, G., Powell, J., Yang, L., Lennon
 G., Hillier, L., Allen, M., Bowles, L., Gelsel, S., Kucaba, T., Marra
 M., Martin, J., Steptoe, M., Tan, F., Theising, B., Bowers, Y., Wylie
 T., Waterston, R., Wilson, R. and Francomano, C.
 WashU-MGB/NHGRI EST Project

TITLE Unpublished (1997)
 JOURNAL Contact: Libin Jia
 COMMENT Medical Genetics Branch
 National Human Genome Research Institute
 10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
 Tel: 301-402-4877
 Fax: 301-496-7157
 Email: libin@helix.nih.gov
 Seq primer: M13 Reverse.

FEATURES
 source Location/Qualifiers

1..354
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="HBMSFLB4"
 /clone_1id="Human Bone Marrow Stromal Fibroblast"
 /sex="Male and Female"
 /tissue_type="bone marrow"
 /cell_type="stromal fibroblast"
 /dev_stage="mixed"
 /lab_host="XLI-blue"
 /note="Vector: pBluescript, Site_1: EcoRI; Site_2: XhoI"
 BASE COUNT 77 a 95 c 95 t 2 others
 ORIGIN

Query Match 23.0%; Score 304.6; DB 9; Length 354;
 Best Local Similarity 94.9%; Pred. No. 2, 1e-88;
 Matches 333; Conservative 0; Mismatches 16; Indels 2; Gaps 2;

OY 497 TCCTCAGCAGGAGGAGGCTCAATTCCTGACTTACCGGGCAGCAAGCGCGTTTC 556
 DB 351 TCTTCAGCAGGAGGAGGAGGCTCAATTCCTGACTTACCGGGCAGCAAGCGCGTTTC 293
 OY 557 GCGAAGTGAAGCGGCTGCTTAATTTCCGCAATCAAACTTGGCTTAAAGTGAAGAC 616
 DB 292 GCGAAGTGAAGCGGCTGCTTAATTTCCGCAATCAAACTTGGCTTAAAGTGAAGAC 234
 OY 617 AGGAGCAAGGCTGCTTAATTCAGGAGGATTAACCATGGAAGTGAAGCGCGGACA 676
 DB 233 AGGAGCAAGGCTGCTTAATTCAGGAGGATTAACCATGGAAGTGAAGCGCGGACA 174
 OY 677 ATGTCATTAACCGGCTGCGGTAGCCCTGCGATCAATGCTGAGGAGATATTTCTCTCC 736
 DB 173 ATGTCATTAACCGGCTGCGGTAGCCCTGCGATCAATGCTGAGGAGATATTTCTCTCC 114
 OY 737 AACAGAGCAGGAGATGCGGAGCGGGGTGGGAGAGATACCGATTACACCACTGGA 796
 DB 113 AACAGAGCAGGAGATGCGGAGCGGGGTGGGAGAGATACCGATTACACCACTGGA 54
 OY 797 ACACCTGCTAAGTTTGATACACCGGCAATTTATTCTCAACGCGAGGCA 849

DB 53 ACACCTGCTAAGTTTGATACACCGGCAATTTATTCTCAACGCGAGGCA 1

RESULT 6 320 bp mRNA 11near EST 18-MAY-2001
 BE436403/c
 LOCUS EST407481 tomato breaker fruit, TIGR Lycopersicon esculentum CDNA
 DEFINITION clone BE436403, mRNA sequence.
 ACCESSION BE436403
 VERSION BE436403
 KEYWORDS GI:9434246
 SOURCE EST.
 ORGANISM Lycopersicon esculentum
 tomato.
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.

REFERENCE 1 (bases 1 to 320)
 Alcala, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
 Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.L., Rønning, C.M.,
 Nielsen, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
 S.D.

TITLE Generation of ESTs from tomato fruit tissue, breaker stage
 JOURNAL Unpublished (2000)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 source Location/Qualifiers

1..320
 /organism="Lycopersicon esculentum"
 /cultivar="T496"
 /db_xref="taxon:4081"
 /clone="CLEG32C20"
 /clone_1id="tomato breaker fruit, TIGR"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCudaapt; Site_1: EcoRI;
 Site_2: XhoI; Fruit were harvested at the breaker stage
 (first sign of lycopene accumulation on the blossom end
 of the fruit). Fruit were cut in half and the seeds and
 locules were discarded prior to freezing the pericarp."
 BASE COUNT 63 a 96 c 86 g 75 t
 ORIGIN

Query Match 16.4%; Score 216.8; DB 10; Length 320;
 Best Local Similarity 95.1%; Pred. No. 1, 4e-59;
 Matches 289; Conservative 0; Mismatches 7; Indels 8; Gaps 6;

OY 242 GACATTACCAAGCGGCTGTGTACCGGAGATTCGTGGGAAAAAGGCTGCCGC 301
 DB 320 GACATTACCAAGCGGCTGTGTGTACCGGAGATTCGTGGGAAAAAGGCTGCCGC 261
 OY 302 AGCTGGCAGGTGGCGATTATTCGTGTCGACGAGCG--TACCGGTAAACAGCGGA 359
 DB 260 AGCTGGCAGGTGGCGATTATTCGTGTCGACGAGCGGTTACCGGTAAACAGCGGA 201
 OY 260 AGCTGGCAGGTGGCGATTATTCGTGTCGACGAGCGGTTACCGGTAAACAGCGGA 141
 DB 360 AGCTGGCAGGTGGCGATTATTCGTGTCGACGAGCGGTTACCGGTAAACAGCGGA 419
 OY 200 AGCTGGCAGGTGGCGATTATTCGTGTCGACGAGCGGTTACCGGTAAACAGCGGA 141
 DB 420 G--TCCAGTCCGATCGTTATTTAA--TCCCTTAAAACTGGC--GTTTCCAACTGGATA 475
 OY 140 GGTCCAGGTCCCGATCCGTTATTTAAATTCCTTAAAACTGGCGTTTGGCCAACTGGATA 81
 DB 476 ACAGGAGCTACGACGAGCGGA--TCCCTGAGAGGGGAGG--AGGGTCAATTTGCTGACTTAC 533
 OY 80 ACAGGAGCTACGACGAGCTTCTTCAGAGGGGAGGAGGTCATTTGCTGACTTAC 21

QY 534 CGGG 537
 Db 20 CGGG 17

RESULT 7
 BM409016/c
 LOCUS
 DEFINITION 208 bp mRNA linear EST 22-JAN-2002
 EST583343 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLE66H24 5' end, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM409016
 EST583343
 CLE66H24
 BM409016
 EST.
 Lycopersicon esculentum
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 208)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
 J., Bougri,O., Kirkness,E., Uterback,T., Van Aken,S., Ronning
 C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 Unpublished (2002)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3

FEATURES
 source

1. 208
 /location="Qualifiers
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLE66H24"
 /clone_1lb="tomato breaker fruit"
 /tissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: BluescriptSKmCunadap; Site.1: EcoRI;
 Site.2: XhoI; supplier: Boyce Thompson Institute; Fruit
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 lycopen accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."
 43 a 56 c 62 g 47 t

BASE COUNT
 ORIGIN
 Query Match 15.7%; Score 208; DB 13; Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4e-57;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 301
 |||||||
 Db 208 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 149
 QY 302 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 361
 |||||||
 Db 148 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 89
 QY 362 CCTTCCGCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGGAGATACGT 421
 |||||||
 Db 88 CCTTCCGCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGGAGATACGT 29
 QY 422 CCAGTCCGATCCGTTATTTAATCCTCT 449
 |||||||
 Db 28 CCAGTCCGATCCGTTATTTAATCCTCT 1

RESULT 8
 BM412806/c
 LOCUS
 DEFINITION 208 bp mRNA linear EST 22-JAN-2002
 EST587133 tomato breaker fruit Lycopersicon esculentum cDNA clone
 CLE66I013 5' end, mRNA sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM412806
 EST587133
 CLE66I013
 BM412806
 EST.
 Lycopersicon esculentum
 Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 208)

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
 J., Bougri,O., Kirkness,E., Uterback,T., Van Aken,S., Ronning
 C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
 Generation of ESTs from tomato fruit tissue, breaker stage (2002)
 Unpublished (2002)
 Contact: CUGI
 Clemson University Genomics Institute
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: http://www.genome.clemson.edu/orders/index.html
 This clone is available through the Clemson University Genomics
 Institute
 Seq primer: T3

FEATURES
 source

1. 208
 /location="Qualifiers
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLE66I013"
 /clone_1lb="tomato breaker fruit"
 /tissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: BluescriptSKmCunadap; Site.1: EcoRI;
 Site.2: XhoI; supplier: Boyce Thompson Institute;
 sequencing: The Institute for Genomic Research. Fruit
 were harvested at the breaker stage (first sign of
 lycopen accumulation on the blossom end of fruit). Fruit
 were cut in half and the seeds and locules were discarded
 prior to freezing the pericarp."
 43 a 56 c 62 g 47 t

BASE COUNT
 ORIGIN
 Query Match 15.7%; Score 208; DB 13; Length 208;
 Best Local Similarity 100.0%; Pred. No. 8.4e-57;
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 242 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 301
 |||||||
 Db 208 GACATACCAACGCCAGCGTCTGTAGCCGACGATTCGTGGGAAAAAGGCTGCCCGC 149
 QY 302 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 361
 |||||||
 Db 148 AGCTGTGTCAGGTGCGGATTAATGTCGATGTCAGCGAGCTACCCGTAACAGCGGAAG 89
 QY 362 CCTTCCGCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGGAGATACGT 421
 |||||||
 Db 88 CCTTCCGCGCGGGCTGGACCTGACCTGTCATTAACCGTACATACCCAGGAGATACGT 29
 QY 422 CCAGTCCGATCCGTTATTTAATCCTCT 449
 |||||||
 Db 28 CCAGTCCGATCCGTTATTTAATCCTCT 1
 RESULT 9
 BJ040796
 LOCUS
 BJ040796 546 bp mRNA linear EST 06-DEC-2001

DEFINITION				
ACCESSION	BX040796 NIBB Mochii normalized Xenopus neurula library Xenopus laevis cDNA clone XL047108 5', mRNA sequence.			
VERSION	BX040796			
KEYWORDS	EST.			
ORGANISM	African clawed frog.			
SOURCE	Xenopus laevis			
REFERENCE	Amphibia: Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia: Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus.			
AUTHORS	1 (bases 1 to 546) Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara,Y.			
TITLE	Expressed genes in X. laevis embryo			
JOURNAL	Unpublished (2001)			
COMMENT	Contact: Tadasu Shin-I Center for Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855 Email: tshin@genes.nig.ac.jp.			
FEATURES	location/qualifiers			
SOURCE	1..546			
	/organism="Xenopus laevis"			
	/db_xref="taxon:8355"			
	/clone_xref="X1047108"			
	/clone_11b="NIBB Mochii normalized Xenopus neurula library"			
	/tissue.type="whole embryo"			
	/dev_stage="stage 15"			
	/note="vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute). "			
BASE COUNT	116 a	145 c	142 g	138 t
ORIGIN	5 others			
Query Match	15.0%; Score 198.8; DB 13; Length 546;			
Best Local Similarity	96.7%; Pred. No. 1.5e-53;			
Matches	203; Conservative 0; Mismatches 7; Indels 0; Gaps 0;			
OY	404 ATACCGACGAGATGTCGTCAGTCGTCGATCCGATCGGTATTTATTCCTTAATAACTGGCGTTT 463			
DB	337 ATTAACTGGCGGTGACGTCGATCCGATCCGTTATTTAACTCTTAATAACTGGCGTTT 396			
OY	464 GCCAACTGGATTAACCGCGAAGCTGATGTCAGCGGATCTTCAGCAGGCGCAGAGGTCATATG 523			
DB	397 GCCAACTGGATTAACCGCGAAGCTGATGTCAGCGGATCTTCAGCAGGCGCAGAGGTCATATG 456			
OY	524 CTGACTTTACCGGGGATCGGCAAAAGGGGTTTCGGCACTGGAACGGGTGCTTAATTTTC 583			
DB	457 CTGACTTTACCGGGGATCGGCAAAAGGGGTTTCGGCACTGGAACGGGTGCTTAATTTTC 516			
OY	584 CGCAATCAAACTTGTGCTTAAACGTGAGA 613			
DB	517 CGCAATCAAACTTGTGCTTAAACGTGAGA 546			
RESULT 10				
BE461872/c	179 bp mRNA linear EST 18-MAY-2001			
LOCUS	EST413391 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA			
DEFINITION	clone CLEG40G22, mRNA sequence.			
ACCESSION	BE461872			
VERSION	BE461872.1 GI:9506174			
KEYWORDS	EST.			
SOURCE	tomato.			
ORGANISM	Lycopersicon esculentum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;			

REFERENCE AUTHORS	TITLE JOURNAL COMMENT	FEATURES source
Lycopersicon. 1 (bases 1 to 179) Alcala,J., Vrebllov,J., White,R., van der Hoeven,R.S., Holt,I.E., Tiang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romling,C.M., Niernan,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley ,S.D.	Generation of ESTs from tomato fruit tissue, breaker stage Unpublished (2000) Contact: CUGI Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Email: http://www.genome.clemson.edu/orders/index.html 5 prime sequence:	Location/Qualifiers 1..179 /organism="Lycopersicon esculentum" /cultivar="T496" /db_xref="taxon:4081" /clone="CLEG40G22" /clone_1db="tomato breaker fruit, TIGR" /tissue_type="Pericarp" /tissue="stage-breaker" /dev_stage="SOLR" /note="Vector: BluescriptSKmCuvadpt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT ORIGIN	40 a 47 c 51 g 41 t	
Query Match Best Local Similarity 96.6%; Score 169.4; DB 10; Length 179; Matches 173; Conservative 0; Mismatches 6; Indels 0; Gaps 0;		
OY	398 CCGTACATATCCAGGACAGATACGTCACGATCCGATCGTTATTTATTCCTTAATAACTG 457	
Db	179 CCGTACATATCCAGGACAGATACGTCACGATCCGATCGTTATTTATTCCTTAATAACTG 120	
OY	458 GCGTTGCCAACTGGATATACCGGACGATGACGATCGATCGTCTGACGAGGAGAGGGT 517	
Db	119 GCGTTGCCAACTGGATATACCGGACGATGACGATCGTCTGACGAGGAGAGGGT 60	
OY	518 CAATTGCTGACTTTCACCGGCGATCGGCAACGCGCTTTCGGAACTGGAACGGGTCTT 576	
Db	59 CAATTGCTGACTTTCACCGGCGATCGGCAACGCGCTTTCGGAACTGGAACGGGTCTT 1	
RESULT 11		
BJ030203	508 bp mRNA linear EST 05-DEC-2001	
LOCUS	BJ030203 N18B Mochli normalized Xenopus neurola library Xenopus	
DEFINITION	laevis cDNA clone X1003H05 5', mRNA sequence.	
ACCESSION	BJ030203	
VERSION	BJ030203.1 GI:17371629	
KEYWORDS	EST.	
SOURCE	African clawed frog.	
ORGANISM	Xenopus laevis	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; Xenopodinae; Xenopus. 1 (bases 1 to 508) Kitleyana,A., Terasaka,C., Mochli,M., Ueno,N., Shin-I,T. and Kohara ,Y.	
REFERENCE AUTHORS	Expressed genes in X. laevis embryo Unpublished (2001) Contact: Tadasu Shin-I Center For Genetic Resource Information National Institute of Genetics 1111 Yata, Mishima, Shizuoka 411-8540, Japan Tel: 81-559-81-6856 Fax: 81-559-81-6855	

	Query Match	10.88;	Score 143.2;	DB 10;	Length 254;
	Best Local Similarity	94.98;	Pred. No.1.8e-35;		
	Matches 148;	Conservative	0; Mismatches 8;	Indels 0;	Gaps 0;
Oy	1 ATGAACGAGTCTTAATCCATTTTTAACTTCCTTGATTCCGTAAACCGCAACTGCA	60			
Db	96 ATGAAGGAGTCTTAATCCATTTTTAACTTCCTTGATTCCGTAAACCGCAACTGCA	155			
Oy	61 TTGCCTCAGAGTAGACCGGAGCTGAAGTGGTGGTATTGTTCAGTCTCATGGT	120			
Db	156 TTCCTCTCAGATGAGCCGAGCTAAGTGGAAAAGTGGATTGTCAAGTCATAGGT	215			
Oy	121 GTGGTGCTCCAACCAAGGCCAACCAACTGATGCAG	156			
Db	216 GTGGTGCTCTCCAACCAAGGCCAACCAACTGATGCAG	251			

RESULT 14
LOCUS H60613/c 392 bp mRNA linear EST 06-OCT-1995
DEFINITION y144d12.s1 Soares fetal liver spleen INFLS Homo sapiens cDNA clone
IMAGE208151 3' similar to gb|L137071|HUMSCALUE Human scRNA molecule
, transcribed from Alu (RNA): gb:J04970 CARBOXYPEPTIDASE M
PRECURSOR (HUMAN); contains Alu repetitive element;; mRNA sequence.
ACCESSION H60613
VERSION H60613.1 GI:1013445
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Metazoa; Primates; Catarrhini; Homidae; Homo.
REFERENCE 1 (bases 1 to 382)
AUTHORS Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Maira,M., Parsons,J.,
Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevasakis,E., Waterston
R., Williamson,A., Wohldmann,P. and Wilson,R.
The WashU-Merck EST Project
Unpublished (1995)
COMMENT Contact: Wilson RK

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Insert Size: 909
High quality sequence stops: 301
Source: IMAGE Consortium, LLNL
This clone is available royalty-free through LLNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 909 Std Error: 0.00
Seq primer: Promega -21ml3
High quality sequence stop: 301.
Location/Qualifiers
I. .392
FEATURES
source

```

BASE COUNT
ORIGIN
116 a      76 c      91 g      98 t      11 others
/db_xref="GDB:3777282"
/db_xref="taxon:9606"
/clone IMAGE:208151"
/clone lib="Soares fetal liver spleen INFLS"
/sex="male"
/dev_stage="20 week post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/note="Organ: Liver and Spleen; Vector: p773D (Pharmacia
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI
1st strand cDNA was primed with a Pac I - oligo(dT) primer
15' AACGTGAGAGATTATATTAACGATCTTTTCTTTTCTTTT 3',
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fátima Bonaldo."
```

	Query Match	3.3%	Score 43.8	DB 14	Length 392
	Best Local Similarity	51.7%	Pred. No. 0.01		
	Matches	93	Conservative	0	Mismatches 87
					Indels 0
					Gaps 0
QY	519	AATTGCTGACTTAAACGGGCGATCGGCAACGGCGCTTTCGCGA	CTCGAACGGGCTGTAA	578	
DB	262	AAGTCTGGGATTCACAGGCGATGAGCCACACNGCCTGGCCATAT	TGGCCTAATGTTTTTAA	203	
QY	579	TTTTTCGGAATCAACACTGTGCCTTAAACGTGAGAAACAGACGA	AGCAAGCTGTTCATTAAC	638	
DB	202	TTATTCACCAAAATTTTGGCCNTTTCTGTGGGAAACAAGTGA	AGCTCTCTTCAGC	143	
QY	639	GCAGCATTAACACTCGAAGCTCAAGGTGAGCGCCGACAAATGT	CTCATTAACCGGTGGGT	698	
DB	142	ATAGACATACCTTATCCCATCATTTTAAATAAACAAGAGTTGT	TTCACCTTCGAGGT	83	

RESULT	15
B1152280/c	
LOCUS	
DEFINITION	B1152280 602917706P1 NCI_CGAP_Lu29 Mus musculus cDNA clone IMAGE:5068177 5'..733 bp mRNA linear EST 05-JUL-2001
ACCESSION	mRNA sequence.
VERSION	B1152280
KEYWORDS	B1152280.1 GI:14612281
SOURCE	EST.
ORGANISM	house mouse.
	Mus musculus
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 733)
AUTHORS	NIH-MGC http://mhc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgaabps-r@mail.nih.gov

CDNA Library Preparation: life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLM11183 row: 1 column: 02
 High quality sequence start: 2
 High quality sequence stop: 602.
 Location/Qualifiers
 FEATURES
 source
 1..733

[illegible]

Db 131 CGGCCCATGATGCTGTGGCGAGTCAACCGGGGCACTGATCATCTGGCTGTG 72
 QY 244 CATTAACACGCGAGGCTGTGTAGCCGACGAGATTGCTG 282
 Db 71 AGGTCCGACGTGAGCGGCTTCTGAGTTCAACAGAGCTG 33

Search completed: June 12, 2003, 12:09:00
 Job time : 2128 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 13:01:29 ; Search time 3587 Seconds

(Without alignments)
3569.901 Million cell updates/sec

Title: US-09-777-566A-2

Perfect score: 2302
Sequence: 1 MKAILPFLSLIPLPQSA.....IVNEARIPACSLSHHHHHH 440

Scoring table: BIOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-rlh
-Q/cgna_1/USPRO.spool/US09777566/runatc_05062003_154808_6550/app.query.fasta.1.583
-DB-GenDb1 -QWt-fastlap -SUFFIX-rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-D1osum62 -TRANS-human40.cdi -LIST=45
-DOCALLIGN=200 -THR_SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT-pto -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09777566 @CGN 1.1 2496 @runatc_05062003_154808_6550 -NCPU=6 -ICPU=3
-NO_MMAP -LARGESQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rnd:*
36: em_htg_man:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	6	AR108133 Sequence
2	2302	100.0	1323	6	AR127818 Sequence
3	2302	100.0	1323	6	AR130956 Sequence
4	2302	100.0	1323	6	AX052713 Sequence
5	2302	100.0	1323	6	AX338966 Sequence
6	2302	100.0	1323	6	AX356566 Sequence
7	2235	97.1	1901	6	EC0APPAA Sequence
8	2235	97.1	1901	6	AX356572 Sequence
9	2235	97.1	3470	6	AX042376 Sequence
10	2235	97.1	4060	6	AX042374 Sequence
11	2235	97.1	6116	6	AX042375 Sequence
12	2235	97.1	6708	6	AX042373 Sequence
13	2235	97.1	10784	6	AE000200 Sequence
14	2235	97.1	17732	6	AX042378 Sequence
15	2235	97.1	18882	6	D90735 Sequence
16	2235	97.1	20623	6	AX042372 Sequence
17	2229	96.8	1901	1	EC0ACPHA Sequence
18	2229	96.8	1901	1	EC0ACPHC Sequence
19	2229	96.8	1901	1	EC0ACPHD Sequence
20	2229	96.8	1901	1	EC0ACPHE Sequence
21	2228	96.8	1901	1	EC0ACPHB Sequence
22	2226	96.7	1901	6	AX356574 Sequence
23	2225	96.7	1901	1	EC0ACPHF Sequence
24	2221	96.5	1901	6	AX356571 Sequence
25	2220	96.4	1901	6	AX356570 Sequence
26	2210	96.0	10029	1	AE005292 Sequence
27	2210	96.0	327773	1	AP002554 Sequence
28	2205	95.8	5421	6	AX042377 Sequence
29	1102.5	47.9	212936	2	AC020970 Sequence
30	943.5	41.0	10029	1	AE013783 Sequence
31	943.5	41.0	193050	1	AJ414149 Sequence
32	570	24.8	114979	2	AC020798 Sequence
33	545	23.7	10278	1	AE005294 Sequence
34	541	23.5	1810	1	PRACPPA Sequence
35	539	23.4	1675	1	EC0AGPA Sequence
36	539	23.4	11710	1	AE000202 Sequence
37	539	23.4	18252	1	D90737 Sequence
38	529	23.0	21670	1	AE008748 Sequence
39	527	22.9	700	1	ECAPPA Sequence
40	526	22.8	10545	1	AE005727 Sequence
41	524	22.8	234050	1	AL627269 Sequence
42	505	21.9	10385	1	AE011713 Sequence
43	505	21.9	13498	1	AE012175 Sequence
44	483.5	21.0	7129	1	AE012063 Sequence
45	373	16.2	11877	1	AE011683 Sequence

RESULT 1

ALIGNMENTS

[illegible]

QY	221	lysvalseralaaaspasvalserleuhrgrlyalavalserleualsexerleuhr	240
Db	661	AAAGGAGGCGGACAAATGTCATTAACCGGGTGGGTAAGCTCGCATCAATGGCTGAGG	720
QY	241	gluilepheleuleuenglnlnalaglnlmetprogluprogl1tYp9lyargrilethr	260
Db	721	GAGATATTTCCTCTCCACAAAGCACAGGGAATGCCGAGCGGGGTGGGGAAGATCACG	780
QY	261	aspserh1sglnttYpasnthrleualseuSerleuH1saana1aglnPheTyrleuLeuGln	280
Db	781	GATTCCACACAGTGGAGACACCTTCTAAGTTTGTCATTAACGGCCAAATTTATTTGCTACAA	840
QY	281	ArgThrProGluValAlaArgSerArg1aThrProleuleuAspLeuileuMet1a1a	300
Db	841	CGCACGCGAGAGTGGCCCGCAGCGCCGACCCGCTTATGGAATTTGATCTGACAGCG	900
QY	301	LeuThrProH1sProGlnProGlnlysgln1a1tYrGlyVal1ThrLeuProH1sValleu	320
Db	901	TTGAGCGCCCATCCACCCCAAAAGCGGTATGGTGTGACATTAACCCACTTCACTACG	960
QY	321	Phe1le1a1glnH1sAspThrAsnleuAlaasaleuGlyGlyAlaLeuGlnleuAsnThr	340
Db	961	TTTATATCCCGGACAGATACTAATCTGGCAATCTCGCGCGGCACTGGAGCTCAACTGG	1020
QY	341	ThrLeuProGlnlyGlnProAspAsnThrProProGlyGlyGlyleuVal1PheGluTrpTr	360
Db	1021	ACGCTTCCCGGTCACCCGGATTAACAGCCGCGCAGGTGGTAAGTGTGTTGTAACGCTGG	1080
QY	361	ArgArgLeuSerAspAsnSerGlnTrp1leGlnVal1SerleuVal1PheGlnThrLeuGln	380
Db	1081	CGTGGGTAAAGGGAATACAGCCAGTGGATTCAGTTTGGTGGTCTTCCAGACTTTACAG	1140
QY	381	GlnMetArgAspLysThrProLeuSerleuAsnThrProProGlyGlyleuVal1LysLeuThr	400
Db	1141	CAGATCGCTGATAAACCCCGCTGCATTAATTAACGCCGCGGAGAGTGAACCTGACC	1200
QY	401	LeuAlaGlyCysGlnGluArgAsn1aGlnGlyMetCysSerleuAlaGlyPheThrGln	420
Db	1201	CTGGCAGAGTGTGAAGAGCGGAATCGCAGGGCATGTCTTCTTGGCAGGTTTACGCCAA	1260
QY	421	1levalaangluAlaArg1leProAlaCysSerleuArgSerH1sH1sH1sH1s	440
Db	1261	ATCGGATGAAGACAGCGCATACCGCGTGCATTTGAGATCTCATCACATCAGCATGCAC	1320
RESULT 2			
ARI27818		1323 bp	DNA
LOCUS	ARI27818		linear
DEFINITION	Sequence 1 from patent US 6183740.		PAT 16-MAY-2001
ACCESSION	ARI27818		
VERSION	ARI27818.1		GI:14115480
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	1 (bases 1 to 1323)		
AUTHORS	Short,J.M. and Kretz,K.A.		
TITLE	Recombinant bacterial phytases and uses thereof		
JOURNAL	Patent: US 6183740-A 1 06-FEB-2001;		
FEATURES	Location/Qualifiers		
source	1..1323		
BASE COUNT	323 a 353 c 357 g 289 t	1 others	
ORIGIN			
Alignment Scores:			
Pred. No.:	3..276-169	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	6	Gaps:	0

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCCTTAACCCCGAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGTGTGATGTCATGTCGTCATGCT 120
 QY 41 ValAlaGlnProThrLeuAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGCTGTCTCCAAACCAAGCCAGCACTGATGAGGATGTCAACCCAGACCACTGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 DB 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCCGNGGTGGTGGACTAATGCCCTATCTC 240
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 241 GGACATTAACCAAGCCAGCTGTGAGCCGAGATTCCTGGCGAAAAGGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGGTGCGCATTAATGCTGATGCGAGACCCGTAACCCAGCGCAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCGCGCGGCTGGCTGACCTGCTGCTGCAATACCGTACATACCCAGCAATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCACTCCCGATCCGTTATTTATCTCTAAAAACCTGGGGTGTGGCACTGGATTAACGG 480
 QY 161 AsnValIleThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGAGTACGACGATCTCTGACGAGGAGGAGGTCAATGTGACTTACCGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAAAAGCGGCTTTCGCGACTGGAACGGGTGCTTAATTTCCGCAATCAATCTTGTC 600
 QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAAGCTGAGAAACAGACGAAAGCTGTTCATTACGACGACATTAACCATCGGACATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGAGCGCGCAATGTCTCATTAACCGGTGCGGTAAAGCTGCAATCAATGCTGACG 720
 QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpProGlyArgIleThr 260
 DB 721 CAGATATTCTCTCTCAACACACAGGAAATCCGCGAGCCGGGGGGGAAAGATCACCC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
 DB 781 GATTCACACACAGTGAACACCTTGTGAAGTTTGCAATACGGCCATTTTATTTGCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCGCAGAGGTTGCCCGCAGCGCGCACCCGCTTATTGATTTGATTCATGCGACGC 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValIleu 320
 DB 901 TTGACGCCCATCTCCACCGAATAACAGGGGTATGGTGTGACATTTACCCACTTCAGTAC 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlnLeuAsnTrp 340
 DB 961 TTTATTTGCGGACGAGATTAATCTGGCAATCTCGGGGGGCACTGGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgThr 360
 DB 1021 ACGCTTCCCGGTACACCGGATTAACAGCGCGCCAGGTGTGAACTGCTTTTGAACGCTGG 1080

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTAAGCGATTAACAGCCAGTGAATCAGTTTGCTGCTGCTTCCAGACTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 DB 1141 CAGATCGCTGATTAACACCGCGCTGATTAATTAATACCGCGCCGAGAGTGAATGACAC 1200
 QY 401 LeuAlaGlyCysGluGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGCGAGATGTGAAGAGCGGAATGCGAGGGCATGTGTCCTTGGCAGGTTTACGCA 1260
 QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 DB 1261 AHCGTGAATGAAGCAGCATACCGCGCTGCACTTTGAGATCTCATCACCATACCATCAC 1320
 RESULT 3
 ARI30956
 LOCUS ARI30956 1323 bp DNA linear PAT 16-MAY-2001
 DEFINITION Sequence 1 from patent US 6190897.
 ACCESSION ARI30956
 VERSION ARI30956.1 GI:14119281
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 1323)
 AUTHORS Kretz, K.
 TITLE Phylase
 JOURNAL Patent: US 6190897-A 1 20-FEB-2001;
 FEATURES Location/Qualifiers
 source 1..1323
 /organism="unknown"
 ORIGIN
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 Alignment Scores:
 Pred. No.: 3, 27e-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566A-2 (1-440) x ARI30956 (1-1323)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGAGGATCTTAATCCATTTTATCTCTTCGATTCCTTAACCCCGAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 61 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGTGTGATGTCATGTCGTCATGCT 120
 QY 41 ValAlaGlnProThrLeuAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGCTGTCTCCAAACCAAGCCAGCACTGATGAGGATGTCAACCCAGACCACTGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
 DB 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCCGNGGTGGTGGACTAATGCCCTATCTC 240
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 241 GGACATTAACCAAGCCAGCTGTGAGCCGAGATTCCTGGCGAAAAGGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGGTGCGCATTAATGCTGATGCGAGACCCGTAACCCAGCGCAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCGCGGCTGGCTGACCTGCTGCTGCAATACCGTACATACCCAGCAATACG 420

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAAATCTCTAAATAACGCGTTGGCACTGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGAGCTGACGCGATCTCAGACGAGGCGGATTAATGCTTACCTTAACTCCGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnAspPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAACGCGGTTCCGGAACTGGAACGGGTCCTTAATTTCCGCAATCAACTGTGCG 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACCTGAAACAGACGAGCAAAAGCTTCTAATACGACGACTTACCATCGGAACTC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACCGCCGACAAATGCTCATTAACCGGTGCGTAAGCTTCCCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyLysArgIleThr 260
 DB 721 GAGATATTCTCTCTCACAACAGACAGGAAATGCCGAGCGGGGTGGGGAAGATCAC 780
 QY 261 AspSerHisGlnThrPheAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 781 GATTCACACAGATGAGACACCTTCTTAAGTTGCTTAACGCGCAATTTATTTGCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIle 300
 DB 841 CGCAGCCGAGAGTTGCCCGACGCGCCGACCCGCTTATTCGATTCATCCTGACAGC 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaLysGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCCATCCACCCCAAAACAGCGCATAGGTGACATTAACCACTTCACTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnThr 340
 DB 961 TTTATTCGCGGACACGATTAATCTGCAAAATCTGCGGCGGCTGAGCTCAACCTG 1020
 QY 341 ThrLeuProGlyLysProAspAsnThrProProGlyGlyLysLeuValPheGlnArgThr 360
 DB 1021 ACGCTCCCGGTCAACCGCGGATTAACGCGCCGAGTGTGAACTGTGTTGAACGCTG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTGGCTTAACGCAATTAACAGCCAGTGCATGAGTTCTGCTGCTCCAGACTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyLysValLysLeuThr 400
 DB 1141 CAGATGCGTGAATAAACCGCCCTGCTAATTAATACGCGCGGAGAGGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnLysMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGAGATGTAAGACGAAATGCGCAGGCGCATGTTCTGCTGGCAGGTTTTCGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
 DB 1261 ATCGTGAATGAAGACGATACGCGCGTGCAGTTGAGATCTCATCACCATCACCATCAC 1320
 RESULT 4
 AX052713 LOCUS AX052713 1323 bp DNA linear PAT 12-JAN-2001
 DEFINITION Sequence 1 from Patent WO0071728.
 ACCESSION AX052713
 VERSION AX052713.1 GI:12226903
 KEYWORDS
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (bases 1 to 1323)

AUTHORS Short, J.M. and Kretz, K.A.
 TITLE Recombinant bacterial phytases and uses thereof
 JOURNAL Patent: WO 0071728-A 1 30-NOV-2000;
 Diverisa Corporation (US)
 FEATURES
 source location/Qualifiers
 1..1323
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1..1323
 CDS
 /note="unnamed protein product"
 /codon_start=1
 /transl_table=11
 /protein_id="CAC21810.1"
 /translation="MKAILIPFLSLILPLPQSAFAQSEPELKESVTVYSRHGVAP
 TKATQIMQDTPDAMPVTPVKGLMTPRGGLIAYLGHYQRLVADGLAKGCPQS
 GOVALIADVDERTKRTGEAFAGLAPDCAITVHQADTSSPDLPENPLKGVCLDPA
 NTPDAIISRAGSITADPFGHOTAEPRELERLNPQSLCKRREKQESGTLNAPS
 ELKYSADVNSLTGAVSLASMLTEFLFDQAQKPEPGVGRITDTSQNTLSLNAPR
 YLQRTPEVARSRAITPPLDLITMAALTTPHPPOKAYGVLTPTSVLEIAGHDTNLNAG
 ALEINWTLPGQPDNTPPGELVFERWRLSDNSQMIQVSLVFTQLOQRDKTPLSLNT
 PPGEVKLTILACSEERNAQMCISLAFQIYVNEARIPACSLSRSHHHHH"
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,27e-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX052713 (1-1323)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAAGCATCTTAATCCATTTTATCTCTTCGATTCGTTAACCCCGCAATCTCA 60
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 61 TTCGCTGACAGTACCGCGAGCTGAAAGTGGAGATGTCAGTCGTCATGCT 120
 QY 41 ValAlaGlnProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGCTGCTCCAAACAGCCAGCCAGCACTGATGAGATGTCACCCCGACGACAGCGCA 180
 QY 61 ThrTrpProValLysLeuGlyLysLeuThrProAspGlyGlyLysGlnLeuIleAlaThrLeu 80
 DB 181 ACCTGGCGGCTGAACCTGGGTGGCTGACACCGGNGGTGGTGAAGTATGCGCTATGTC 240
 QY 81 GlnHisThrGlnArgGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
 DB 241 GGACATTAACCAACGCGACGCTGCTGACCGACGATGCTGGCGAATAAGGCTGCCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyLys 120
 DB 301 CAGCTGTGTCAGGTCGCGCATTAATGCTGATGTCGCGAGCGCTACCCCGTAAACAGCGCA 360
 QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGCGCGCGGCTGGCACCTGACTGTGCATTAACCGTACATCAACGAGCATGAC 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCGTTATTTAAATCCTCTAAATAACGCGTTGGCAACTGATTAACGCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGAGCTGACGCGATCTCAGACGAGGCGGATTAATGCTTACCTTAACTCCGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnLysArgValLeuAsnAspPheProGlnSerAsnLeuGly 200

Db 541 CGGCAAAAGCGGCTTTCGGAACTGGAGACGGGTGCTTAATTTCCGCAATCAAACTTGTC 600
 QY Leuysaragluylusglnaspglusercysserleuthrglnalaaleuprosergluenu 220
 Db 601 CTTAAAGCTGAGAAACAGAGAGAAAGCTGTTCAATTAACCGAGCATTTACCATCGAATCTC 660
 QY LysvalseralaaspasnvalserleuthrglyalavalserleualaSermetleuthr 240
 Db 661 AAGGTAGCGCCGCAACATGCTCATTAACGGGTGGGTAAGCCTCGCATCATGCTGACG 720
 QY Glutlepheluleuenglnalaglnlglmetprogluproglytyrpglyargllethr 260
 Db 721 GAGATATTTCCTCCGCAACAGACAGGAAATGCCGGAGCCGGGGTGGGAAGGATCAC 780
 QY Asperthsglntppasnthleuenserleuhsasnalaglnphtyrleuengln 280
 Db 781 GATTCACACCACTGGACACCTTGCTAAGTTGCAATACGGCAATTTATTTGCTACAA 840
 QY Argthrprogluvalalaargserarglaathrproleuaspheullemetalaala 300
 Db 841 CGCAGCCGAGAGGTGGCCGAGCGCGCCACCCGCTTATTGGATTGATCATGCGACG 900
 QY Leuthrprohispptroglinlyslnalatryglyvalthrleuprothrservalneu 320
 Db 901 TTGACGCCCCCATCCACCCGCAAAACAGCGATAGTGTGACATTAACCATTCACACTG 960
 QY Pheilealaglyhsaspthrasnleualaenleuglyglalaleuenglenuenrpp 340
 Db 961 TTATATGCGGACAGATACTAATCTGCAAAATCTGGCGGCGCACTGAGCTCAACTGG 1020
 QY Thrleuproglylnproaspasnthrproglglygluleuvalphegluartrpp 360
 Db 1021 ACGCTCCCGGTCAGCGGATTAACACGCGCGGAGGTGTAACCTGTTTAAAGCGCG 1080
 QY Argargleuseraspasnserglntrpilleglinalserleuvalpheglintrleu 380
 Db 1081 CGTCGGCTTAACGATTAACAGCCAGAGTTCAGGTTTCGCTGCTCCACACTTAAACG 1140
 QY Glnmetargaspbylthrproleuserleuasnthrproglglygluleuvallysleu 400
 Db 1141 CAGATGCGGTAAACCCGCTGCTAATTAATACGCGCGCGGAGGTGAACATGACC 1200
 QY Leuualaglycysgluaglnargasnalainglmetcysserleualaglyphenr 420
 Db 1201 CTGGCAGATGGAAGAGCAAGAAATGCGAGGCAATGCTGCTGGCAGGTTTACGCA 1260
 QY Ilevalasnglualarglleproalacysserleuarqserhishishishis 440
 Db 1261 ATCGTGAATGAAGCAGCATACCGGCGTGCAGTTTGAGATCTCATCACCATCATCAC 1320
 RESULT 5
 LOCUS AX338966 1323 bp DNA linear PAT 09-JAN-2002
 DEFINITION Sequence 1 from Patent WO0189317.
 ACCESSION AX338966
 VERSION AX338966.1 GI:18129102
 KEYWORDS
 SOURCE Escherichia coli.
 ORGANISM Escherichia coli.
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
 REFERENCE 1
 AUTHORS Short, J.M., Kreitz, K.A. and O'Donoghue, E.
 TITLE Dietary aids and methods of use thereof.
 JOURNAL Patent: WO 0189317-A 1 29-NOV-2001;
 DIVERSA CORPORATION (US)
 FEATURES
 source
 1. 1323
 /organism="Escherichia coli"
 /db_xref="taxon:562"
 1. 1323
 /note="unnamed protein product"
 /codon_start=1

/transl_table=1
 /protein_id="CAD20553.1"
 /db_xref="gi:18129103"
 /translation="MKATLPELILPLPPOSAPFAOSEPELKIESVYISRRHGVRAP
 TKATQMDVTPDAMPPTTPPVKLTGMLTPRGSELAVLHYQADTSPPDLPLPKTVCQDINA
 GOVALIADVDERTKRTGEAFAGLAPDCAIPDCAITVHROADTSPPDLPLPKTVCQDINA
 NWTDALISRRAGSLADPEFHROTAFRELERLANPPOSMLCLKRKODESCSLTALPS
 ELKYSADVSLTGAVSLASMLTEIFLQAOAGMPEGMGRITSDSHOWMTLSLNAOF
 YLORTPEVARSRAPIPLDLIMALTLPPOKQAYGVTLPTSVYLEIGHDTNLIAG
 ALRNPTLRGQPDNTPRGSELVFERMRSLSDNSQIMQVSLVFTQLQWRDTPISLNT
 PPGVKRLTLACCEERNAGKCSLAFGTQIVNEARIPKCSLSHHHHH"
 BASE COUNT 323 a 353 c 357 g 289 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 3 276-169 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0
 US-09-777-566a-2 (1-440) x AX338966 (1-1323)
 QY 1 Metlysalaileleuilepropheluserleuileproleuthrproglinserala 20
 Db 1 ATGAAGGAGNCTTAATCCATTTTATCTCTGATTCGTTAACCCGCAATCTGCA 60
 QY 21 phealaglnsergluproglyleuylsleuengluservalvallevalserarghls 40
 Db 61 TTCGCTCAGAGTACCCGAGCTGAAGCTGGAAGCTGCGATGTGTCAGTCATCGT 120
 QY 41 Valargalaprothrlthylalathrglnleuemetglnasppvalthrproaspala 60
 Db 121 GTCGCTGCTCCACCAAGGCCACGCACTGATGAGATGTCACCCAGAGCATGGCCA 180
 QY 61 Thrtpptrovallysleuglytyrpleuthrproargglygluleuilealatrleu 80
 Db 181 ACCTGCGCGTAAACCTGGGTGGCTGACACCGCGNGGTGGTACCTAATGCGCTATCTC 240
 QY 81 Glyhistryglinarglargleuvalalaaspglyleuvalalyslysglycyspro 100
 Db 241 GGACATTTACCAAGCGCACGCTGCTAGCCGCGATTCGCGCAAAAGCGCTGCCG 300
 QY 101 Glnserglylnvalalietllealaspvalaaspgluaargthrarglysthr 120
 Db 301 CAGTGTGTCAGTGCAGATTAATCTGATGTGCAAGAGCGTAAACCGTAAACAGCGA 360
 QY 121 Alaphaalalaglyleualaproaspysalalietlnvalhsthrghinalaaspphr 140
 Db 361 GCCTTGGCGCGCGGCTGGCACCTGACTGTGCAATACCGTACATACCCAGGCAATACG 420
 QY 141 SerSerProaspptroleupheasnproleuylsthrglyvalcysglnleuaspasna 160
 Db 421 TCCAGTCCCGATCCCTTAATTAATCTCTAAAAATGCGCGTTCGCCAATGATTAACCG 480
 QY 161 Asnvalthrapsalilileuserargylaglyysertllealasphehrglyhis 180
 Db 481 AACGTAGTACGACGCAATCTTCAGCAGGCGAGGAGGCTCAATGCTGATCTTAACCGGAT 540
 QY 181 ArgglnthrAlaphaarggluleuaglualargvalleuanpheproglinserasnleucys 200
 Db 541 CGGCAAAAGCGGCTTTCGGCAATGGAACGGGTCTTAATTTCCGCAATCAAACTTGTCG 600
 QY 201 Leuylsaragluylusglnaspglusercysserleuthrglnalaaleuprosergluenu 220
 Db 601 CTTAAAGCTGAGAAACAGAGCAAGCTGTTCAATTAACGGGCAATTCATCGGAATCTC 660
 QY 221 LysvalseralaaspasnvalserleuthrglyalavalserleualaSermetleuthr 240
 Db 661 AAGGTAGCGCCGCAACATGCTCATTAACGGGTGGGTAAGCCTCGCATCATGCTGACG 720
 QY 241 Glutlepheluleuenglnalaglnlglmetprogluproglytyrpglyargllethr 260

```

|||||
Db 721 GAGATATTTCCTCCCAACAGCAGGGAATCCGGAGCCGGGGTGGGGAAGATCAC 780
Oy 261 AspSerHisGlnThrPasnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGAGTGAACACCTTGTGAAGTTGGCTAAACGCCAATTTATTTGCTACAA 840
Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAla 300
Db 841 CGCAGCCGAGAGGTGGCCGAGCCGCCACCCCGTTATTGGATTGATCATGACAGC 900
Oy 301 LeuThrProHisProGluGlnValArgValAlaThrLeuProThrSerValIleu 320
Db 901 TTGACGCCCATCCACCCCAAAACAGCGATGATGATGACATTACCCACTTCAGTACTG 960
Oy 321 PheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnThr 340
Db 961 TTTATTTGGCGGACACGATACACTGATCTGGCAAAATCTCGCGCGGCGACATGAGTCAAC 1020
Oy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgThr 360
Db 1021 ACGCTTCCCGGTACCCGAGTAAACAGCCGCCAGGTGGTGAACCTGCTTTGAACGCTCG 1080
Oy 361 ArgArgLeuSerAspAsnSerGlnThrPrlGlnValSerLeuValPheGlnThrLeuGln 380
Db 1081 CGTCGGCTAAGGATTAACAGCCAGTGATTCAGGTTGCTGCTTCCAGACTTACAG 1140
Oy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1141 CAGATGCGTGATAAACCGCGCTGTCAATTAATACGCCGCCGAGAGGTGAACATGAC 1200
Oy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGAGATGGAAGACGGAATCGCAGGCGATGCTGCTGCGAGGTTTATCGCAA 1260
Oy 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGAATGAAGACGATACCGCGGTGACGTTTGAGATCTCATACCATCATCATCAC 1320

RESULT 6
AX356566 1323 bp DNA linear PAT 06-FEB-2002
LOCUS AX356566
DEFINITION Sequence 1 from Patent WO0190333.
ACCESSION AX356566
VERSION AX356566.1 GI:18621053
KEYWORDS
SOURCE
ORGANISM Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Escherichia.
REFERENCE
1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and
O'Donoghue, E.
Recombinant bacterial phytases and uses thereof
JOURNAL Patent: WO 0190333-A 1 29-NOV-2001;
DIVERSA CORPORATION (US)
FEATURES
SOURCE
1. 1323
/location/Qualifiers
/organism="Escherichia coli"
/db_xref="taxon:562"
1. 1323
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAD23025.1"
/db_xref="GI:18621053"
/translation="MKATILFPLSLILPVTQSAFAOSEPEIKLESVIVSRHGVAP
TKATQLMDVTPDAMPVPRKLGWLTTPRGELIAYLGXORQRLVADLAKGCPQS
GOVAIIADVDERTRKTEGEAFAGLAIDELIVHQADSSDPLEPKTGVCOLDNA
NVTDAIISRAGSIAEDFGHROTAFRELEVLNFPNSNLCLKREODSGSLTOAPS
ELKYSADNVSLTGAIVSLASMLTEIFLQOAGMPGKRGITDSHOWMTLISNAOC
YLIDRTPEVASRATPLILDITMALITPPPOKAYGVTLPTVSLEFIAGHDYNLANLG
ALBNWTLPGOPDNTPEGELVFEKWRKLSNSQMIQVSLVFLOQMRDXTPLSLNT

```

```

BASE COUNT 323 a 353 c 357 g 289 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 3,276-169 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0
DB: 6
US-09-777-566a-2 (1-440) x AX356566 (1-1323)
Oy 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGGATCTTAATCCATTTTATCTGTGATTCGGTTAAACCCCAATCTGCA 60
Oy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTGAACCCGAGCTGAAGCTGGAAGTGGTGATTTGATGTCATGCTATG 120
Oy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 121 GTGCTGCTCCACCAAGGCCACGCACTGATGAGATGTCACCCCAAGCATGCGCA 180
Oy 61 ThrTrpProValLysLeuGlyTyrPleuThrProArgLysGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGCGCGGTAAACTGGGTGGCTGACACCGCGNGGTGGAGCTAATCGCATCTC 240
Oy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GCACATTACCAAGCCAGCAGCTCTGTGTACCGCAGATTGCTGGCCAAAAGGGCTGCCG 300
Oy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTGTGCTCAGTGCCTTATTTCTGATGTCGACGAGCTACCCCGTAAACAGCGCAA 360
Oy 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCCTTCGCGCGCGGCTGGACACTGATGTGCAATTAACCGTACATACCAGACATACG 420
Oy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCCTTATTAATCCTTAATAAATGCGCTTGGCACTGATTAACCG 480
Oy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCTACGACGCGCAGAGGCTCAATTCGTGACTTTACCGGCA 540
Oy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db 541 CGGCAACGCGGCTTCCGGAACCTGGAACGGGGCTTAATTTCCCAATCAACACTTGTC 600
Oy 201 LeuLysArgGluGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACGTGAAGAACAGAGCAAAAGCTTCAATTACGACGAGCTTCATCGGAATC 660
Oy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACGCCGCAATGCTCTATTAAACCGGTGGTGAAGCTGCAATGCTGACG 720
Oy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglYArgIleThr 260
Db 721 GAGATATTTCCTCCCAACAGCAGGGAATGCCGAGCCGGGGTGGGGAAGATCAC 780
Oy 261 AspSerHisGlnThrPasnThrLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTACACACGAGTGAACACCTTGTGAAGTTGGCTAAACGCCAATTTATTTGCTACAA 840
Oy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAla 300
Db 841 CGCAGCCGAGAGGTGGCCGAGCCGCCACCCCGTTATTGGATTGATCATGACAGC 900

```

OY		301	LeuP ^{HprOHI} SProProGlnInLySlaIatHygLYaIThrleuproThrSerValIeu	320
Dd		901	TTGAGCCCCATCCACCGCAAAACAGCGGTATGGTGTGAATTACCCTTCAGTACTG	960
OY		321	PheIIeaIagLYHISAspPT ^H AsnIeuValaSnIeuGlyGLYAlaIeUGlIleuAsnTrp	340
Dd		961	TTTATATGCCGAGACAGATACTAATCTGGCAAATCTCGCGCGCATGAGACTCAACTGG	1020
OY		341	ThrIeuProGIyGlnProAspnThrProFroGIyGIUleuValIPhegluarTrp	360
Dd		1021	ACGCTTCCCggTGACCGGATTAACAGCCGGCAGSTGGTGAACGTGGTGTGAACGCATGG	1080
OY		361	ArgarGlueSerAspAsnSerGlnTrpIleGlnValSerIeuValIPheGlnThrIeuGln	380
Dd		1081	CGTCGGCTAACAGATACAGCACGCGAGGTTCAGGTTCGCTGGCTTCACACTTTACAG	1140
OY		381	GlnMetArqAspLYst ^H ThrProIeuSerIeuAsnTrpProFroGIyGluValIylValSer	400
Dd		1141	CAGATGCTGTATAAAACCGCCGTGCATTAATATACGCCGCCGAGAGAGTGAACCTGACC	1200
OY		401	LeuAlaGIcYSgSLugLUARgaSnaIagInGIymetCysserIeuAlaGlp ^H ethrGln	420
Dd		1201	CTGGCAGGATGTGAAGACGAAATGCCAGGGCANGTTCCTTGCGCAGGTTTACGAA	1260
OY		421	IlevalaSnGUaluaIaArgIleproAlaCysserIeuArgSerHisHisHisHis	440
Dd		1261	ATCGGATGATGAAGCAGCATACCGCGCTGAGATTGAGATGTCATCACCATCACCATCAC	1320
RESULT 7				
LOCUS	ECOAPPAA	1901 bp	DNA	linear BCF 04-APR-2002
DEFINITION	Escherichia coli periplasmic phosphoanhydride phosphohydrolase (appa) gene, complete cds.			
ACCESSION	M58708			
VERSION	M58708.1	GI:145283		
KEYWORDS				
SOURCE				
ORGANISM	Escherichia coli. Escherichia coli. Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.			
REFERENCE	1 (bases 1 to 1901) Dassa,J., Marcq,C. and Boquet,P.L. The complete nucleotide sequence of the Escherichia coli gene appa reveals significant homology between pH 2.5 acid phosphatase and glucose-1-phosphatase J. Bacteriol. 172 (9), 5497-5500 (1990)			
JOURNAL				
MEDLINE	90368616			
PUBMED	2168385			
REFERENCE	2 (sites) Ostanin,K., Harms,E.H., Stevis,P.E., Kuciel,R., Zhou,M.M. and Van Etten,R.L. Overexpression, site-directed mutagenesis, and mechanism of Escherichia coli acid phosphatase J. Biol. Chem. 267 (32), 22830-22836 (1992)			
AUTHORS				
TITLE	3 (sites) Ostanin,K. and Van Etten,R.L. Asp04 of Escherichia coli acid phosphatase is involved in leaving group protonation J. Biol. Chem. 268 (28), 20778-20784 (1993)			
JOURNAL				
MEDLINE	94012612			
PUBMED	8407904			
FEATURES				
source	Location/Qualifiers 1..1901 /organism="Escherichia coli" /db_xref="taxon:562" 3..1522 /gene="appa" 3..8 /gene="appa" 16..108			
gene				
RBS				
CDS				

stem_loop	/gene="appa" /note="Orfx" /codon_start=1 /transl_table=11 /product="unknown" /protein_id="AA072085.1" /db_xref="GI:145284" /translation="MWYLLMFVGIILMCSTLVVWLDRLKS" 118. .149
-35_signal	/gene="appa" 133. .138
-10_signal	/gene="appa" 156. .161
RBS	/gene="appa" 172. .177
CDS	/gene="appa" 188. .1486 /standard_name="acid phosphatase" /codon_start=1 /transl_table=11 /product="periplasmic phosphoanhydride phosphohydrolase" /protein_id="AA072086.1" /db_xref="GI:145285" /translation="MKAILIPPLSLILPITPQSAFAOSEPELKIESVIYSRRGVRAP TKATQLDVYTPDAMPVTPVVKIGWLTTPRGELIAYLDHYQORLVAAGLAKKCCPDS GCVAILADVDETRKTGEAFAGLAPDCALIVTHQADTSSPDPLFNLTAKGVCLDNA NTVDLISRAGSIIADFTGHROTAFRELERYLNFPOSMLCKREKODESCLTADPS ELKASADVNLGVAISLASMLTEIFILQOAGMPDPGMRITDSHOKNTLISLHNOF YILQTPVARSRRATPLDLIKRTALTPTPPPKOKAYGTYLPTSYLFIGHPTNLANLGG ALIELWTLTPGCDNTPGGGELVFERRWRLSDNSQIOWSLVFTQLQOMRKPTPLSLNT PGEVKKTLTGCEERNAGMCSLAGFTQIVNEARIPACSL" 188. .253
sig_peptide	/gene="appa" 254. .1483
mat_peptide	/gene="appa" /product="periplasmic phosphoanhydride phosphohydrolase" 299. .301
variation	/gene="appa" /standard_name="H303A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R16A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R20A mutant" /note="created by site directed mutagenesis" /citation=[2] 299. .301
variation	/gene="appa" /standard_name="R92A mutant" /note="created by site directed mutagenesis" /citation=[2] 302. .304
variation	/gene="appa" /standard_name="H17N mutant" /note="created by site directed mutagenesis" /citation=[2] 1163. .1165
variation	/gene="appa" /standard_name="D304A mutant" /note="created by site directed mutagenesis" /phenotype="Asp 304 changed to Ala" /replace="gct" 1163. .1165
variation	/gene="appa" /standard_name="D304E mutant" /note="created by site directed mutagenesis" 1163. .1165

/citation-[3]
/phenotype="Asp 304 changed to Glu"
/replac="gag"
1497. .1522
/gene="appa"
/note="rho-independent terminator"

BASE COUNT 474 a 499 c 499 g 428 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 7,71e-164 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 1 Gaps: 0

US-09-777-566a-2 (1-440) x ECOAPPAA (1-1901)

OY 1 Metlysalaileuleilepropheluseuleuleileproleuthrproglinserala 20
DB 188 ATGAAGAGCATCTTAATCCATTTTATCTTCGATTCGTTAACCCGCAATCTGCA 247
OY 21 Phealaginsergluuprogluuleulysleuulservalillevalserarghisgly 40
DB 248 TTCCCTCAGAGTGAGCCGAGCTGAAGCTGGAAGTGCTGATGTCAGTCGTCATGCT 307
OY 41 Valargalaprotlhrlysalatthrcgluulemetglinservalthrproaspalatp 60
DB 308 GTGGGTGCTCCACCAAGCCAGCCAGCACTGATGCGAGTGCACCCGACGCAATCGCA 367
OY 61 Thrtrprovalylsleuglytyrpleuthrproarglyglyluuleilealatyrl 80
DB 368 ACCGTGCGCGTAAACTGGTGTGCTGACACCGCGGTGCTGAGCTATGCGCTATCTC 427
OY 81 GLYHSTYRGINARGGLINARGLEUVALAASPGLYLEUVALALYSGLIYCYSP 100
DB 428 GGACATTACCAACGCGAGCTGTGTTAGCCGAGATTGCTGGCAAAAAGGCGTCCCG 487
OY 101 Glnserglygluvalaileilelelaaspvalaspgluargthrarglythr 120
DB 488 CAGTCTGTGTCAGTCCGATTTATTCCTGATGTCGACGACCGTACCCGTAACAGCGCA 547
OY 121 Alaphealaialaglyleualaiproaspysalailethrvalhsthrglnaaspthr 140
DB 548 GCCTTCGCGCGCGGCTGCGACCTGACTGCAATTAACCCATACCCAGGCAATACG 607
OY 141 Sersepproasprouleupheasnproleuylsthrglyvalcyglnleuaspasna 160
DB 608 TCCAGTCCCGATCCGTTATTTAACTCTTAATAAACTGGCTTGGCAATCGATTAACCG 667
OY 161 Asnvalthrasspalailleuseraraglaiglyserilealaasphethr 180
DB 668 AACGTGATGACGCGATCTCAGCAGGCGAGGAGGTCAATTCGCACTTACCGCGCAT 727
OY 181 Arglnthrailaphearglualeugluarvalleuasnphetroglinsersanleucys 200
DB 728 CGGCAAAACGGGTTCCGCAACTGGAAGCGGTGCTTATTTCCGCAATCAAACTTGTGC 787
OY 201 Leuysarggluylusglnasppluserysserleuthrlnalaleuprosersglu 220
DB 788 CTTAAACGTGAAACAGGAGAAAGCTGTTCAATTAAAGCAGCATTAACATCGGAACTC 847
OY 221 Lysvalseralaaspasnaivalserleuthrcllyalavalserleualaasermetleuthr 240
DB 848 AAGGTGACCGCGCAATGTCTCATTAACCGGTGGGTGACCTGCTCATGCTGACG 907
OY 241 Gluilepheleuleuglnalaglnlymetprogluproglytyrtplyargllethr 260
DB 908 GAGATATTTCTTCGCAACAGCAGGAGATGCGGAGCGGGGTGGGAAGATCAC 967
OY 261 Asperthrsgintpaspthrleuleuserleuhsanlaaglnpheryrleuleugln 280
|||||

DB 968 GATTCACACAGTGGACACCTTGCTTAAGTTTGCATTAACGCCCAATTTTATTCCTACAA 1027
OY 281 Argthrprogluvalaialargseraraglatthrproleuileaspillemetalaia 300
DB 1028 CGCAGCGCAGAGTGGCCCGCAGCGCGCCACCCGTTATTAAGATTGATCAAGACAGC 1087
OY 301 Leuthrprohispoproglinlysclinalatyrglyvalthrleuprothrserval 320
DB 1088 TTGACGCCCATCCACGCCAAAACAGCGCTATGTGTGACATTAACCATTCAGTGTCTG 1147
OY 321 Pheilealaglyhsthraspthrasnleuilaasnleuglylalaileuglualeuasn 340
DB 1148 TTTATCGCGGACACGATTAATCTGCAATTCGCGCGCGCTGAGCTGAGCTACACTG 1207
OY 341 Thrleuproglyglinproaspasnthrproproglygluileuvalphegluargtrp 360
DB 1208 ACGTTCGCCGTCACCCGATTAACCGCGCGCAGGTGTGAATGTGTGTAACCGCTGG 1267
OY 361 Argargleuseraspasnserglintpilleglinalserleuvalpheglinthrleugln 380
DB 1268 CGTGGCTTAACGATTAACAGCAGTGTGATTCAGTGTGCTGCTTCACAGCTTTACAG 1327
OY 381 Glnmetargasplysthrproleuserleuasnthrproproglygluvallysleuthr 400
DB 1328 CAGATCGCTGAATAAACCCCGCTGTCAATTAATACCGCGCGGAGAGGTGAACGTAGC 1387
OY 401 Leualaglycygsluarguargasnalainglymetcysserleualaaglypethr 420
DB 1388 CTGGCAGATGTGAAAGACGAATATCGCAGGCAATGTCTGTGGCAGGTTTACGCA 1447
OY 421 Ilevalaengluvalaargyllerproalacysserleu 432
DB 1448 ATCGTAATGAAGCAGCATACCGCGCTGCGAGTTTG 1483
RESULT 8
AX356572 1901 bp DNA linear PAT 06-FEB-2002
LOCUS AX356572
DEFINITION Sequence 7 from Patent WO0190333.
ACCESSION AX356572
VERSION AX356572.1 GI:18621059
KEYWORDS
SOURCE
ORGANISM
Escherichia coli.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.
REFERENCE
1 Short, J.M., Kretz, K.A., Gray, K.A., Barton, N.R., Garrett, J.B. and O'Donoghue, E.
TITLE Recombinant bacterial phytases and uses thereof
JOURNAL Patent: WO 0190333-A 7 29-NOV-2001;
DIVERSA CORPORATION (US)
FEATURES
source
1. 1901
Location/Qualifiers
/organism="Escherichia coli"
/db_xref="taxon:562"
BASE COUNT 474 a 499 c 499 g 428 t 1 others
ORIGIN
Alignment Scores:
Pred. No.: 7,71e-164 Length: 1901
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 6 Gaps: 0
US-09-777-566a-2 (1-440) x AX356572 (1-1901)
OY 1 Metlysalaileuleilepropheluseuleuleileproleuthrproglinserala 20
DB 188 ATGAAGAGCATCTTAATCCATTTTATCTTCGATTCGTTAACCCGCAATCTGCA 247
OY 21 Phealaginsergluuprogluuleulysleuulservalillevalserarghisgly 40
|||||

```

Db      248  TTCCTCAGACTGAGCCGAGCTGAAGCTGGAAGTGTGATGTGTCAGTCGTATGCT 307
Oy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      308  GTGGGTGCTCCAAACAGCCAGCCAGCACTGATGCGAGGTGTCACCCAGCAGCAGGCCA 367
Oy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db      368  ACCTGGCCGGAATAAAGTGGTGTGGCTGACCCGGGNGTGGTGAAGCTAATGCCCTATCTC 427
Oy      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      428  GGACATTAACCAACCCAGCCGCTGTGATGCCAGCATGTCGGCAAAAAGGGCTGCCCG 487
Oy      101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgGlyThrArgLysThrGly 120
Db      488  CAGTCTGTGTCAGTGGCGCATTAATGCTGATGTGCGAGCGCTACCCGTAACACAGCGCA 547
Oy      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      548  GCCTTCGCGCGCGGCTGGCAGCTGCTGCAATACCGTACATACCCAGCAGCATACG 607
Oy      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      608  TCCAGTCCCGGATCCGTTATTAATCCCTAAAACTGGCGTTGCCCACTGGATACGCG 667
Oy      161  AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      668  AACGTGACGTACGGGATCTCAGCAGAGGCGAGGCTCAATGTGCTGACCTTACCGGGAT 727
Oy      181  ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnProGlnSerAsnLeuGly 200
Db      728  CGGCAACAGCGGTTTCGGAACCTGAGACGGGTGCTTAATTTCCCAATCAACTGTGTC 787
Oy      201  LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      788  CTTAAAGCTGAGAAACAGAGCAAAAGCTGTTCAATTAACGACGATACATCGGAACTC 847
Oy      221  LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      848  AAGGTGAGCGCGCAATGTCTCATTAACCGGTGGGTGAACCTCTGCATCAATGTGTGAG 907
Oy      241  GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr 260
Db      908  GAGATATTTCCTCGCAACACAGCAGGATGCCGAGGCCGGGTGGGAAGATCAACC 967
Oy      261  AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db      968  GATTTCACACAGGTGGAACACCTTGCTTAAGTTTGCAATACGCGCAATTTTATTTGCTACA 1027
Oy      281  ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db      1028  CGCAGCGCAGAGGTGGCCGCGAGCGCCGCGGCTTAATTAAGATTATTAACAAGCACGC 1087
Oy      301  LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db      1088  TTGAAGCGCCCATCCACCGCAAAACAGCGCATGGTGTGACATTAACCCACTTCAGTGTG 1147
Oy      321  PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      1148  TTTATTCGCGCAGACGATACATACTGCAATCTCGCGCGCGCACTGAGACTCAACTGCG 1207
Oy      341  ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1208  ACCTTCCTCCGGTCCAGCGGATTAACACGCGCGCAGGTGTGTAAGTGTGTAAGCGTGG 1267
Oy      361  ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1268  CGTGGGCGTAACGATTAACAGCAGTGTGATTCAGGTTTGTGCTTCCAGACTTTAAG 1327
Oy      381  GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400

```

```

Db      1388  CAGATGCTGATTAACACCGCTGTCAATTAATACCCCGCGAGAGGTGAACCTGACC 1387
Oy      401  LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      1388  CTGGCAGATGTGAAGACGGAATGCGAGGGCATGTGTGTTGCGAGGTTTACGCA 1447
Oy      421  IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      1448  ATCGTGAATGAAGCAGCATACCGCGTGCAGTTTG 1483

RESULT 9
AX042376          3470 bp   DNA       linear   PAT 23-NOV-2000
LOCUS             Sequence 5 from Patent WO0064247.
DEFINITION        AX042376
ACCESSION         AX042376.1 GI:11340994
VERSION           AX042376.1
KEYWORDS           synthetic construct.
SOURCE             synthetic construct.
ORGANISM           artificial sequences.
REFERENCE          1 (bases 1 to 3470)
AUTHORS            Forsberg,C.W., Golovan,S. and Phillips,J.P.
TITLE              Transgenic animals expressing salivary proteins
JOURNAL            Patent: WO 0064247-A 5 02-NOV-2000;
                   UNIVERSITY OF GUELPH (CA)
FEATURES
   source           1. 3470
                   /organism="synthetic construct"
                   /db_xref="taxon:32630"
                   /note="R15/APPA transgene"
BASE COUNT         1065 a      721 c      735 g      949 t
ORIGIN
Alignment Scores:
Pred. No.:         1,57e-163      Length:      3470
Score:             2235.00         Matches:      430
Percent Similarity: 99.54%         Conservative: 0
Best Local Similarity: 99.54%      Mismatches:  2
Query Match:       97.09%          Indels:       0
DB:                6              Gaps:          0

US-09-777-566a-2 (1-440) x AX042376 (1-3470)
Oy      1  MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db      1811  ATGAAGCATCTTAATCCATTTTATCTCTTGATTCGCTTAACCCGCAATCTGCA 1870
Oy      21  PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db      1871  TTCGCTCAGAGTACCGGAGCTGAAGCTGGAAGTGTGTGATGTGTCAGTCGTATGCT 1930
Oy      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      1931  GTGCTGTCTCCAAACCAAGCCAGCACTGATGCGAGTGTCCACCCAGACGATGGCCA 1990
Oy      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db      1991  ACCTGGCCGGAATAAAGTGGTGTGGTACACCGCGGTGTGTGAGCTAATGCCCTATCTC 2050
Oy      81  GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db      2051  GGACATTAACCAACGCGCAGCTGTGTAGCCGAGGATTTGCGCAAAAAGGGCTGCCCG 2110
Oy      101  GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgGlyThrArgLysThrGly 120
Db      2111  CAGTCTGTGTCAGTGGCGCATTAATGCTGATGTGCGAGCGCTACCCGTAACACAGCGAA 2170
Oy      121  AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      2171  GCCTTCGCGCGCGGCTGGCAGCTGCTGCAATTAACCGTACATACCCAGCAGATACG 2230
Oy      141  SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160

```


Db 2231 TCCAGTCCCGATCCGTTATTATCTCTAAAAAATGGCGCTTGGCAACTGATACCGG 2290
 Qy 161 AsnValThrAspAlaIleuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 2291 AACGACTGACGCGATCTCGACAGGAGGAGGTCATATGCTGACTTTACCGGGCAT 2350
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluLysValIleuAsnPheProGlnSerAsnLeuGly 200
 Db 2351 CGGCAACGGCGCTTCCGGAACGTGACGGGTCTTAATTTTCCCAATCAAACTTGTC 2410
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAAACGTGAAACAGACGAAAGCTGTCATTACAGCAGGCACTTACATCGGAATC 2470
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTAGACCGCGCAATATGTCATTAAACGGGTGGGTGATTAACATGCTGACG 2530
 Qy 241 GluIlePheLeuLeuGlnAlaGlnAlaGlnMetProGluProGluProGluArgIleThr 260
 Db 2531 GAGATATTTCCTCCGCAACAGCACAGGGAATGCCGAGCCGGGGTGGGAAGATCAC 2590
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 2591 GATTACACCAAGTGAACACCTGCTAAGTTGCTAAACGCCCAATTTATTGCTACAA 2650
 Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
 Db 2651 CGCAGCGCAGAGGTGGCCGCGCGCCGACCCGCTTATGATTTGATCAAGACAGG 2710
 Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 Db 2711 TTGACGCCCCATCCACCCCAAAACAGCGCATAGGTGACATTACCCACTTCACGTGCG 2770
 Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnThr 340
 Db 2771 TTTATCGCCGACACGATACCTATGCGCAATCTCGCGCGGCGCATCTGAGCTCACTG 2830
 Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgThr 360
 Db 2831 ACGCTTCCCGGTCAACCCGATACACGCCCGCAGGTGTAACCTGCTTGAACGCTGG 2890
 Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 2891 CGTGGCGTCAACGATACAGCAGGATGATGAGTTGCTGCTTCACAGCTTAAAG 2950
 Qy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValIleLysLeuThr 400
 Db 2951 CAGATGCGTGAATMAACGCCCTGTCATTAAATACGCCGCCGAGAGAGTGAACTGAC 3010
 Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlnMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGAGAGATGTGAAGACGAAATCGCAGGAGCATGTCGTTGCGCAGGTTTACGCA 3070
 Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGTAATGAAGCACGATACCGCTTGCAGGTTTG 3106
 RESULT 10
 AX042374 4060 bp DNA linear PAT 23-NOV-2000
 LOCUS AX042374
 DEFINITION Sequence 3 from Patent WO0064247.
 ACCESSION AX042374
 VERSION AX042374.1 GI:11340992
 KEYWORDS
 ORGANISM
 SOURCE
 synthetic construct.
 artificial sequences.
 REFERENCE
 1 (bases 1 to 4060)
 Forsberg, C.W., Golovan, S. and Phillips, J.P.
 Transgenic animals expressing salivary proteins
 Patent: WO 0064247-A 3 02-NOV-2000;
 JOURNAL UNIVERSITY OF GUELPH (CA)

FEATURES
 source 1..4060 Location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="R15/APPA + Intron transgene"
 BASE COUNT 1257 a 814 c 843 g 1146 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,89e-163 Length: 4060
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: Gaps: 0
 US-09-777-566a-2 (1-440) x AX042374 (1-4060)
 Qy 1 MetLysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1811 ATGAAGCCATCTTAATCCATTTTATCTCTGATTCCTTAACCCGCAATCTGCA 1870
 Qy 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 Db 1871 TTGCTCAGAGTGAGCGGAGCTGAAGTGGAAGTGTGTAATGTGTGCTGCTGATGCT 1930
 Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 Db 1931 GTGGCTGCTCCCAACCAAGGCCACGCAATGATGAGAGATGTCAACCCAGACGATGGCA 1990
 Qy 61 ThrTrpProValLysLeuGlyTyrPheLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 Db 1991 ACCCTGCGCGTAAACTGGGTGGCTGACACCGCGCGGTGGAGCTAATCCCTATCTC 2050
 Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 Db 2051 GGACATTCACACGCCGACGCTGTGGTAGCCGACGATTCCTGCCAAAAAGGCGTCCCG 2110
 Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 2111 CAGCTGGTCAGCTGCGCATTTATGCTGATGTCGACGAGCGATACCCGTTAAACAGGCGCA 2170
 Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 2171 GCCTTGCGCGCGGCTGACCTGATGTCGTAATACCGTCAATACCCAGGACAGATACG 2230
 Qy 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 2231 TCCAGTCCCGATCCGTTATTATCTCTTAATAACTGGCGTTTGCCAACTGATACCGG 2290
 Qy 161 AsnValThrAspAlaIleuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 Db 2291 AACGACTGACGCGATCCTCGACAGGAGGAGGTCATATGCTGACTTTACCGGGCAT 2350
 Qy 181 ArgGlnThrAlaPheArgGluLeuGluLysValIleuAsnPheProGlnSerAsnLeuGly 200
 Db 2351 CGGCAACGGCGCTTCCGGAACGTGACGGGTCTTAATTTTCCCAATCAAACTTGTC 2410
 Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAAACGTGAAACAGACGAAAGCTGTCATTACAGCAGGCACTTACATCGGAATC 2470
 Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTAGACCGCGCAATATGTCATTAAACGGGTGGGTGATTAACATGCTGACG 2530
 Qy 241 GluIlePheLeuLeuGlnAlaGlnAlaGlnMetProGluProGluProGluArgIleThr 260
 Db 2531 GAGATATTTCCTCCGCAACAGCACAGGGAATGCCGAGCCGGGGTGGGAAGATCAC 2590
 Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 2591 GATTACACCAAGTGAACACCTGCTAAGTTGCTAATACCGGCAATTTATTGCTACAA 2650

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
|||||
Db 2651 CGCAGCCGACAGGTTGCCCCGACGCCGCCACCCCGTTATGATTTATATCAAGACACG 2710
QY 301 LeuThrProHisProProGluInLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
2711 TTGAGGCCGCCATCCACGCCAAGAGCGTATGGTGTGACATTACCCACCTTCAGTGTCTG 2770
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
2771 TTTATCGCCGACACGATACCTAATCTGCAATCTCGCGCGGCGCACTGAGCTCAACTCG 2830
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAlaArgTrp 360
2831 ACGCTTCCCGGTCCAGCCGATTAACGCCGCCAGGTGTGTGAATGTGTTTGAACGCTGG 2890
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
2891 CGTCGGCTAAGCATTAACAGCCAGTGCATGAGTTTGGTCTTCACACTTACAG 2950
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnTrpProProGlyGlyValLysLeuThr 400
2951 CAGATGCGGTGATAAACGCCGCTGCTCATTAATACGCCGCCGAGAGGTGAACCTGACC 3010
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
3011 CTGGCAGGATGTGAAGACCGAAATGCCAGGGCATGTGTTCCTTGCCAGGTTTTCACCA 3070
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
3071 ATCGTGATGAAGCAGCATACCGCTTGCAGTTTG 3106
Db 3071 ATCGTGATGAAGCAGCATACCGCTTGCAGTTTG 3106

RESULT 11
AX042375 6116 bp DNA linear PAT 23-NOV-2000
LOCUS AX042375
DEFINITION Sequence 4 from Patent WO0064247.
ACCESSION AX042375
VERSION AX042375.1 GI:11340993
KEYWORDS

SOURCE synthetic construct.
ORGANISM synthetic construct.
REFERENCE 1 (bases 1 to 6116)
AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
TITLE Transgenic animals expressing salivary proteins
JOURNAL Patent: WO 0064247-A 4 02-NOV-2000;
UNIVERSITY OF GUELPH (CA)
FEATURES
source location/Qualifiers
1..6116
/organism="synthetic construct"
/db_xref="taxon:32630"
/note="R15/APPA plasmid with pBLCAT3 vector"

BASE COUNT 1724 a 1386 c 1407 g 1599 t
ORIGIN

Alignment Scores:
Pred. No.: 3,06e-163 Length: 6116
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 6 Gaps: 0

US-09-777-566a-2 (1-440) x AX042375 (1-6116)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1811 ATGAAGGCATCTTAATCCCATTTTATCTCTTCGATTCGCTTAACCCGCAATCTGCA 1870
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 1871 TTCGCTCAGAGTGACCGGAGCTGAAGCTGGAAGCTGTGTGATTTGTCAGTGTGATGAT 1930

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
1931 GTGGCTGCTCCACCAACGACGACGACGATGATGACGATGTCACCCGACGACGACGCA 1990
QY 61 ThrTrpProValLysLeuGlyTyrLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
1991 ACGTGGCGGTAAGAACTGGGTGGGTGACACGCGCGGTGGTGAAGCTAATGCGCTATCTC 2050
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
2051 GGACATTTACCAACGCGACGCTGTGTACCCACGATTTGCTGGCGAAGAGGCTGCCG 2110
QY 101 GlnSerGlyGlnValAlaIleIleLeuAspValAspGluArgGlyThrArgLysThrGlyGlu 120
2111 CAGTGTGTGACGTTGCGCATTTATGTGATGTGACAGCGGTACCCGTAACAGCGCA 2170
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
2171 GCCTTGGCGCGCGGCTGGACACTGATGCAATTAACCGTAATACCCAGCAGCATACG 2230
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
2231 TCCAGTCCCGGATCCGTTATTAATCTTAATAAGCTGCGTTGCCAACTGATTAACGCG 2290
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
2231 AACGTCATGACGCGCATCTCAGCAGGCGAGGAGGTCAATTCGCTGACTTACCGCGCAT 2350
QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuGly 200
2351 CGGCAAGCGCGGTTTCGGAACGTGAACGGGGCTTAATTTCCGCAATCAAACTGTGCG 2410
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
2411 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGATTAACCATCGGAATGC 2470
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
2471 AAGGTGACGCGCGACAAATGTCATTAAACCGGTGGGTGAAGCTGCAATGCTGACG 2530
QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPheLysArgIleThr 260
2531 GAGATATTTCCTCTCAACAGACGACAGGAATGCCGAGCGGGGTGGGAAGATATCAC 2590
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
2591 GATTTCACACACAGTGAACACCTTGTGAAGTTGCAATACGCGCAATTTATTTGCTACAA 2650
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
2651 CGCAGCCGACAGGTTGCCCGCAGCGCGCCACCCCGTTATGATTTATGATCAAGACACG 2710
QY 301 LeuThrProHisProProGluInLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
2711 TTGAGGCCGCCATCCACGCCAAGAGCGTATGGTGTGACATTACCCACTTCAGTGTCTG 2770
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
2771 TTTATCGCCGACACGATACCTAATCTGCAATCTCGCGCGGCGCACTGAGCTCAACTCG 2830
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluAlaArgTrp 360
2831 ACGCTTCCCGGTCCAGCCGATTAACGCCGCCAGGTGTGTGAATGTGTTTGAACGCTGG 2890
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
2891 CGTCGGCTAAGCATTAACAGCCAGTGCATGAGTTTGGTCTTCACACTTACAG 2950
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnTrpProProGlyGlyValLysLeuThr 400
2951 CAGATGCGGTGATAAACGCCGCTGCTCATTAATTAACGCCGCCGAGAGGTGAACCTGACC 3010

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGAGAGATGTGAAGACGAAATCGCAGGACGATGTCTGTCGAGGCTTTACGAA 3070
 QY 421 IleValaGlnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGATGAATGAAGACGATACCCGCTTGACGATTG 3106
 RESULT 12
 AX042373 6708 bp DNA linear PAT 23-NOV-2000
 LOCUS Sequence 2 from Patent WO0064247.
 DEFINITION AX042373
 ACCESSION AX042373.1 GI:11340991
 KEYWORDS
 ORGANISM
 SOURCE
 synthetic construct.
 artificial sequences.
 1 (bases 1 to 6708).
 REFERENCE
 AUTHORS Forsberg,C.W., Golovan,S. and Phillips,J.P.
 TITLE Transgenic animals expressing salivary proteins
 JOURNAL Patent: WO 0064247-A 2 02-NOV-2000;
 UNIVERSITY OF GUELPH (CA)
 FEATURES
 source 1. 6708
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="R15/APPA + Intron plasmid with pBICAN3 vector"
 BASE COUNT 1916 a 1479 c 1515 g 1798 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 3.42e-163 Length: 6708
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 Gaps: 0
 US-09-777-566a-2 (1-440) x AX042373 (1-6708)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 Db 1811 ATGAAGCATCTTAATCCCATTTTATCTCTCGATTCGCTTAACCCCGCAATCTCA 1870
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 Db 1871 TTCGCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTCGTCATGGT 1930
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 Db 1931 GTGGCTGCTCCACACACGACGACGACGATGCGAGATGTCACCCCGACGACGATGGCA 1990
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 Db 1991 ACCTGCGCGGTAACACTGGTGGCTGACACCGCGGTGGTGAAGCTAATGCCATCTC 2050
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuAlaLysLysGlyCysPro 100
 Db 2051 GGACATTACCAACCCACGCGCTGTGTACCCACGATGTCGGGAAAAAGGCGTGGCCG 2110
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 Db 2111 CAGTCTGCTCAGGTCGCGCATTTATGCTGATGTGACGAGCTACCCCGTAACAGCGCAA 2170
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 Db 2171 GCCTTCGCGCGCGGCGTGGCAGCTGACGTGTCATTAACCGTACATCCAGGATACG 2230
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 Db 2231 TCCAGTCCCGATCCGTTATTAATCTCTAAAACTGGCGTTTGCACAACTGATTAACCGC 2290

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 Db 2291 AACGTGACTGACGCGATCTTCAGCAGCAGAGAGGATCAATTGCTGACTTACCGCGCAT 2350
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnIleValLeuAsnProGlnSerLysLeu 200
 Db 2351 CGGCAACGCGCTTTCGCAACTGGACGAGGCGTGAATTTTCCCAATCAAACTGTGCG 2410
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 Db 2411 CTTAACGTGAGAAACAGACGAAAGCGTTCATTAAACGACGATTTACCATCGGAACCTC 2470
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 Db 2471 AAGGTGACGCGCGCAATGCTCATTAACCGGTGCGTGAAGCTTCATCATGCTGACG 2530
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyLysArgIleThr 260
 Db 2531 GAGATATTTCTCTCCACACACGAGGAAATGCGGAGCGGGGTGGGAAAGATCACCC 2590
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 Db 2591 GATTCACACACAGTGAACACCTGCTAAGTTGCAATACGCGCATTTTATTTGCTACAA 2650
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 Db 2651 CGCAGCCGAGAGGTGCGCGCAGCCGCGCCACCCCGTTATTAAGATTGATTAACAGACGCG 2710
 QY 301 LeuThrProHisProProGluGlnAlaTyrGlyValThrLeuProThrSerValGlu 320
 Db 2711 TTGACGCCCATCCACCGCAAAACAGCGATGATGATGATTAACCATTCATCACTGCTG 2770
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 Db 2771 TTTATCGCCGCGCAGCATACTAATCTGGCAATCTCGCGCGGCGCATCGAGTCAACTCG 2830
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 Db 2831 ACGCTTCCCGGACCCCGATTAACACGCGCGCATGTGTGAATCTGTGTGAACGCTGCG 2890
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 Db 2891 CCTGCGCTAAGCATTAACGACGAGTGTTCAGTTCAGTTCGCTTCACAGCTTAACG 2950
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 Db 2951 CAGATGCGCTGATAAAGCCCGCTGCTCATTAATACGCGCGCGAGAGTGAACACTGACC 3010
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 Db 3011 CTGGCAGATGTGAAGACGAAATGCGCAGGCGATGTCTGCTGCGCAGGTTTAACGCA 3070
 QY 421 IleValaGlnGluAlaArgIleProAlaCysSerLeu 432
 Db 3071 ATCGATGAATGAAGACGATACCCGCTTGACGATTG 3106
 RESULT 13
 AE000200 10784 bp DNA linear BCT 01-DEC-2000
 LOCUS Escherichia coli K12 MG1655 section 90 of 400 of the complete genome.
 DEFINITION AE000200 U00096
 ACCESSION AE000200.1 GI:2367111
 VERSION AE000200.1 GI:2367111
 KEYWORDS
 SOURCE
 ORGANISM
 Escherichia coli K12.
 Bacteria: Proteobacteria; gamma subdivision: Enterobacteriaceae;
 Escherichia.
 REFERENCE
 AUTHORS Blattner,F.R., Plunkett,G. III, Bloch,C.A., Perna,N.T., Burland,V.,
 Riley,M., Collado-Vides,J., Glasner,J.D., Rode,C.K., Mayhew,G.F.,
 Gregor,J., Davis,N.W., Kirkpatrick,H.A., Goeden,M.A., Rose,D.J.,
 Mau,B. and Shao,Y.

TITLE The complete genome sequence of *Escherichia coli* K-12
 JOURNAL Science 277 (5311), 1453-1474 (1997)
 MEDLINE 97426617
 PUBMED 9278503
 REFERENCE 2 (bases 1 to 10784)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 3 (bases 1 to 10784)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecollegenetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 4 (bases 1 to 10784)
 AUTHORS Plunkett, G. III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 On Sep 9, 1997 this sequence version replaced gi:1787214.
 This sequence was determined by the E. coli Genome Project at the
 University of Wisconsin-Madison (Frederick R. Blattner, director).
 Supported by NIH grants HG00301 and HG01428 (from the Human Genome
 Project and NCHRR). The entire sequence was independently
 determined from E. coli K12 strain MG1655. Predicted open reading
 frames were determined using Genemark software, kindly supplied by
 Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
 30332 [e-mail: mark@ember.gatech.edu]. Open reading frames that
 have been correlated with genetic loci are being annotated with CG
 Site Nos., unique ID nos. for the genes in the E. coli genetic
 Stock Center (CGSC) database at Yale University, kindly supplied by
 Mary Berlyn. A public version of the database is accessible
 (http://cgsc.biology.yale.edu). Annotation of the genome is an
 ongoing task whose goal is to make the genome sequence more useful
 by correlating it with other data. Comments to the authors are
 appreciated. Updated information will be available at the E. coli
 Genome Project's World Wide Web site
 (http://www.genetics.wisc.edu). *** The E. coli K12 sequence and
 its annotations are periodically updated; this is version M54. No
 sequence changes. Annotation updates: updated gene identifications
 and products; all new functional assignments courtesy of Monica
 Riley; added promoters, protein binding sites, and repeated
 sequences described in reference 1. The unique numeric identifiers
 beginning with a lowercase 'b' assigned to each gene (protein- or
 RNA-encoding) are now designated as gene synonyms instead of
 labels. This should allow them to be searched for in Entrez as gene
 names.

FEATURES
 SOURCE Location/Qualifiers
 1. 10784
 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="MG1655"
 /db_xref="taxon:83333"
 49. 84
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 123. 1421
 /gene="appa"
 /note="b0980"
 123. 1421
 /gene="appa"
 /EC_number="3.1.3.2"
 /EC_number="3.1.3.26"
 /function="enzyme: Central intermediary metabolism. Pool,
 multipurpose conversions"
 /note="0432; 100 pct identical to PPA_ECOLI SW: P07102"
 /codon_start=1
 /transl_table=1
 /product="phosphoanhydride phosphorylase; pH 2.5 acid

phosphatase; periplasmic"
 /protein_id="AAC74065.1"
 /db_xref="GI:1787215"
 /translation="MKAILIFPLSLIPLTPQSAFPAQSEPELKLSEVYISRRGVRAP
 TKATQIMQDVTPTMPVVKIWMTPRGELIAYLGHYQQRILVAGLAKKCCPS
 GOVAIADIADVDETRKGTAEAFAGLAPCAITVHQADTSFDPFLNPKTGVCOLDNA
 NVTDAIISRAGSIADEFGHROTFRLEERLAINFPOSNCLIKRQEQSCSLTALPS
 ELKASADNVSLTGAIVSLASMETEIFLQQAQMGPEPGMGRTDSHOMNTLISNAQF
 YLLORTPEVASRATPLDLIKTALTHPHPOKAYGCTLPASVYFIAGHDTNLNLG
 ALLENMTLPQGPDPPTPGGELVEFRMRRLSDNSQWISVLPOTLQMDRDTPLSLNT
 PPGVKRLTLGACCEERNMAQMCSLAGFQIYVEARIIPACSL"
 1153. 1189
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 repeat_region
 1491. 1524
 /note="REP (repetitive extragenic palindromic) element;
 contains 1 REP sequence"
 gene
 complement(11536. 3716)
 /gene="yccc"
 /note="b0981"
 complement(11536. 3716)
 /gene="yccc"
 /function="orf: Unknown"
 /note="f726; 100 pct identical to fragment YCCC_ECOLI
 SW:P38134 but has 627 additional N-terminal aa and 2 less
 C-terminal residues"
 /codon_start=1
 /transl_table=1
 /product="orf, hypothetical protein"
 /protein_id="AAC74066.1"
 /db_xref="GI:1787216"
 /translation="MTTKNMNTPGSTQENIDLRVLGELMDHRKFLISTATFTLI
 AVASLSTPIYQADTLVQVQKGNMISLSTMINSSPESPEIOLLOSLTICK
 TITALENLDIYEQKFTPIVGRKARLIRKREGLAISNMIPLQNGDQQLTGVN
 GHYTLDEEFTVQNVQGVQEKDVALTIADIKAKPGTQFVLSQRTLEIALNQLQET
 TVSSRSKESQMLELTMWGDDPLITRLINSANNYLOONLQAQDSQSELEFQRL
 PVESELDQAEKINLVYRQODSDVLNLEAKAVLEQIYVNDQNLTLFRAEISQY
 KKHPTFALLERKOTLEQERKRLNKRVSAMPSTQOEYLRSDVEGAVYLAAMLR
 QOELISKSSAIGNVRILIDPAVNOPOVYKPKKALNVYLGIFLGISGVATLQAMR
 RGVAPRQLEERHGISYATTPMSERLDRKTRFKRNKIPSNQQRRTNIPFLANDNA
 DSAVEAVALRTSLHFMETENINILITGATPGSGTFVSTLAAVIAQSDQVLT
 DADLRGVSNNLFTVSNHGLSEYLAKDELDNKVIOHFGKGFVIRGVQPPNSEL
 LMRMRQOLEMADHYDLVIVPPMLAVSDAIVGRSVSTSLVARFGLNTAKES
 LSMQRDQAGVINGKIALGVYIKRASAVSYGVYVYSSEK"
 complement(13733. 3764)
 /note="factor Sigma70; predicted +1 start at 1043443"
 complement(13736. 4194)
 /gene="yccy"
 /note="b0982"
 complement(13736. 4194)
 /gene="yccy"
 /function="putative enzyme; Not classified"
 /note="f152; This 152 aa ORF is 33 pct identical (13 gaps)
 to 139 residues of an approx. 160 aa protein PPAI_SCHPO
 SW: P41893"
 /codon_start=1
 /transl_table=1
 /product="putative phosphatase"
 /protein_id="AAC74067.1"
 /db_xref="GI:1787217"
 /translation="MDSNMAQLKFNLSILVYGTGICRSPISGERLLRRLPGVKYSAG
 VHGLVKHPADATADVAANHGVSLEGHGRGLTFPMARNVDLITLAMESEHIAQTATA
 PEVGRKTMLEQGMLEQKEIPPRYKKSQDAFHYVGMLEKRSQENAKRLSR"
 complement(13793. 3820)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043503"
 complement(13818. 3844)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043528"
 complement(13865. 3892)
 /gene="yccy"
 /note="factor Sigma70; predicted +1 start at 1043575"
 complement(14170. 5309)
 gene

CDs

/gene="ycsc2"
/note="b0983"
complement(4170. .5309)
/gene="ycsc2"
/function="phenotype: Not classified"
/note="f379; This 379 aa ORF is 29 pct identical (33 gaps) to 364 residues of an approx. 400 aa protein BEXD_HAEIN
SW: P2236"

/transl_table=1
/codon_start=1
/product="putative function in exopolysaccharide production"
/protein_id="AAC74068.1"
/db_xref="GI:1787218"
/translation="MKKIEFSLVLAIVLSLACTLVPQGLSTSNKDIETLPDNOY LIDKMNIVPTPLIDOLRAKPIMSQANPELEQIANVEYRIGIGDVLATVMDHPE LTPAGQVRSASDQNMVNDGALFPIPTIGLAKAGTLTVQVREITRRLDSVIESRO VDSVAARSOAKATVTESEKSGQPLTINLITMDATNAAGGLTADMDRNVALVTO GVKTRVLYALMQGDLRQNKLLHPGDLFIPRNDLKVFVAGVGKOSTLKMDSGM TLALALNABGMQDVAATGIFVIRATONKGIANTYOLNKRDSAMILGTEFOL EPRDIVVTTPALRMNRVISLAVPTISGVHDLTETSRWIOQTMFN"
complement(5355. .7451)
/gene="ymca"
/note="b0984"
complement(5355. .7451)
/gene="ymca"
/function="orf: Unknown"
/note="f698; This 698 aa ORF is 64 pct identical (7 gaps) to 694 residues of an approx. 704 aa protein Y0BH_ECOLI
SW: P32689"

gene

CDs

/codon_start=1
/transl_table=1
/product="orf, hypothetical protein"
/protein_id="AAC74069.1"
/db_xref="GI:1787219"
/translation="MKKSYLSGLATVASSCAAEVLTYPDPLSGSSDDEGCTGLLO MPNARIPEGEFSVNYRNDQYRFTSTSVAFPMLEGTIRTTDTRTKYSQWEPESGD QSTKDSFDEKLMEGTYLQVAFGRDLGAGLFDGEYLAVSKQAGPDPFTLGMA WGVAGNAGNTINPCRVSDKCYHRAESHADGDISDFRGPASIFGIEVOTPMNPL

Alignment Scores:

Pred. No.:	5,98e-163	Length:	10784
Score:	2235.00	Matches:	430
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	97.09%	Indels:	0
DB:	1	Gaps:	0

US-09-777-566a-2 (1-440) x AE000200 (1-10784)

QY 1 MetLysAlaIleuLeuProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
Db 123 ATGAACGCGATCTAATCCCATTTTATCTCTCGATTCCGTTAACCCCGCAATCTCA 182

QY 21 PheAlaGlnSerGluProGluLeuLeuLeuSerValValIleValSerArgHisGly 40
Db 183 TTCCCTCAGAGTGGCCGAGAGCTGAAGCTGGAAAGTGGGATGTCAGTCGATCATGT 242

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 243 GTGGCTGCTCCACCAAGCCAGCCCACTGATGCAGATGTCACCCCGAGCATGGCCA 302

QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuAlaTyrLeu 80
Db 303 ACCTGGCGGTAAACAGCTGGTGGCTGACACCGCGGTGGTGAAGCTAATCGCTATCTC 362

QY 81 GlyHisArgGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db 363 GGACATTACCAACCCGAGCGCTGCTGATGCCAGCATTTGCTGGGAAAAAGGCTGCCCG 422

QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
Db 423 CAGTCTGCTCAGCTGGCATTTATGCTGATGTCGACAGCAGCTACCCGTAACAGCGCA 482

QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 483 GCCTTCGGCGCGGGCTGGACCTGATGTCGAATTAACCGTACATACCAGACGATACG 542

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 543 TCCAGTCCCGATCCGTTATTTAATCTCTAAACATGGCGTTTGGCAACTGGATTAACCG 602

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
Db 603 AACGTACTGACGCGATCTCAGCAGGAGAGAGGTCAATTTGCTGACTTACCGGCAT 662

QY 181 ArgGlnThrAlaPheArgGlyLeuGlnLuarGlyValLeuAsnPheProGlnSerAsnLeuCys 200
Db 663 CGGCAACGCGGTTTCCGGAACTGGAAGGGGTGCTTAATTTCCGCAATCAAGCTTGTC 722

QY 201 LeuLysArgGlyLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 723 CTTAAACGCTGAGAAACAGCAGCAAGAGCTTTCATTAAACGACGCAATTCGCAACTC 782

QY 221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
Db 783 AAGGTAGCGCGCAATGCTCTATTACCGGTGGTAAAGCTTCGCAATGCTGACG 842

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyTrpGlyArgIleThr 260
Db 843 GAGATATTTCCTCCCAACAGCAGCAGGAAATGCCGAGCGGGGTGGGAAGATCACG 902

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db 903 GATTTCACACAGTGGAAACCTTGCTAAGTTGCTTAACGGCAATTTATTTTGGTACAA 962

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 963 CGCAGCGCAGAGTTCGCCGAGCGCGCCACCCCGTTATTAAGTTATCAAGACAGCG 1022

QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1023 TTGACGCCCATCCACCCCAAAACAGCGCTGATGTCGACATTACCCACTTCAGTGGCTG 1082

QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db 1083 TTTATCGCGGAGACGATCTAATCTGGCAAAATCTGGCGGCGGACCTGGACCTCAACTGG 1142

QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGlnArgTrp 360
Db 1143 ACGCTTCCCGGTCAAGCCGGATTAACAGCGCCGAGGTGTGAACGTGTGTAACGCTGG 1202

QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1203 CGTGGCTTAACGATTAACAGCAGCAGTGATGAGTTTGGTTCCTTCCACACTTTAAG 1262

QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1263 CAGATGCGGTAAACCGCCCTGTCAATTAATTCGCGCGGAGAGGTAAACGTGACC 1322

QY 401 LeuAlaGlyCysGlyGlnLuarGlnAlaGlnLysMetCysSerLeuAlaGlyPheThrGln 420
Db 1323 CTGGCAGGATGTGAAGACGAAATGCGCAGGCGCATGTGTGTTGGCAGGTTTAAACGAA 1382

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1383 ATCGGATGAAGCAGCATACCGCGCTGCAAGTTTG 1418

RESULT 14

AX042378 17732 bp DNA linear PAT 23-NOV-2000
LOCUS
DEFINITION Sequence 7 from Patent W00064247.
ACCESSION AX042378
VERSION AX042378.1 GI:11340996
KEYWORDS
SOURCE synthetic construct.

ORGANISM synthetic construct
 REFERENCE 1 (bases 1 to 17732)
 AUTHORS Forsberg, C.W., Golovan, S. and Phillips, J.P.
 TITLE Transgenic animals expressing salivary proteins
 JOURNAL Patent: WO 0064247-A 7 02-NOV-2000;
 UNIVERSITY OF GUELPH (CA)
 FEATURES
 source 1. 17732
 location/Qualifiers
 /organism="synthetic construct"
 /db_xref="taxon:32630"
 /note="Llama2/APPA transgene"
 BASE COUNT 4719 a 4125 c 4168 g 4719 t 1 others
 ORIGIN
 Alignment Scores:
 Pred. No.: 1.08e-162 Length: 17732
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: Gaps: 0
 US-09-777-566a-2 (1-440) x AX042378 (1-17732)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 12653 ATGAAAGCCATCTTAATCCATTTTATCTCTGATTCGCTTAACCCGCCAATCTGCA 12712
 QY 21 PheAlaGlnSerLupProGlnLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 12713 TTGCTCAGAGTAGGCGGAGCTGAAGCTGAAGTGTGATTCAGTCGATGCT 12772
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 12773 GTGCGTGTCAACCAAGGCGACGACACTGATGATGATGATGATGATGATGATGATGAT 12832
 QY 61 ThrTrpProValLysLeuGlnGlyTrpLeuThrProArgLysGlnLeuIleAlaTyrLeu 80
 DB 12833 ACCGCGCGGTAAACAGCGGTTGGCTGACACCGCGCGTGTGATGATGATGATGATGATGAT 12892
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 12893 GGACATTAACCAAGCGCGCTGTGTGACGCGGATGCTGGCGGAAAGGCGTGGCCG 12952
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspLupArgThrArgLysThrGlyGlu 120
 DB 12953 CAGCTGTGTCAGGTGCGGATTAATGCTGATGTCAGACAGCGGTACCCCTAAACAGCGCAA 13012
 QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 13013 GCCTTCCGCGCGCGCGCTGCGACCTGACTGTGCATTAACCTTAACCTTAACCTTAACCTTA 13072
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 13073 TCCAGTCCCGATCCGTTATTAATCTCTAAACAGCGCGTTCCTCAACGAGTAACGCG 13132
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 13133 AACGTGCTCAGCGACGACCTGACAGCGAGGAGGCTCAATTCCTTACCGGCGCAT 13192
 QY 181 ArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
 DB 13193 CGGCAAAAGCGCGTTTCGCGACCTGGAACGGGTCTTAATTTTCCGCAATCAACCTTGTGC 13252
 QY 201 LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 DB 13253 CTAAACGCTGAGAAACGAGCAAAAGCTGTTCAATTAACGAGCATTAACATGCAAGCTC 13312
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 13313 AAGGTGAGCGCGCAACATGCTCATTAACCGGTGCGGTAAAGCCCTGCATCAATGCTGACG 13372

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrpGlyArgIleThr 260
 DB 13373 GAGATATTTCCTCCGCAACAAACAGACAGGAAATGCGGAGCGCGGTGGGAGAGATACCC 13432
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 DB 13433 GATTCACACAGTGTGAACACCTTCCTGAATTTGATTAACGCGCAATTTTATTCCTCAAA 13492
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 13493 CGCAGCGCAGAGGTTGGCCGACGCGCCACCCGTTATTAATTTGATTAACAGACAGCG 13552
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 13553 TTGACGCCCATCCACCGCAAAACAGCGGATGATGATGATGATGATGATGATGATGATGAT 13612
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 13613 TTTATCCCGGACACGATTAATCTGGCAATCTCGCGCGCGCACCTGCACTCAACTG 13672
 QY 341 ThrLeuProGlnGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrp 360
 DB 13673 ACCTTCCCGGTACGCGGATTAACACCGCCGACGCTGATGATGATGATGATGATGATGAT 13732
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 13733 CGTGGCTAAGCGATTAACAGCACTGATTCAGTTCGCTGCTTCCTCCAGACTTTACAG 13792
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
 DB 13793 CAGATGCGGTGTAACAGCGCGCTCTCATTAATACCGCGCGCGGAGAGGTAACCTGACC 13852
 QY 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGCAGAGTGTGAAGCGGAATGCCAGGCGCATGTGTGTTGCGAGGTTTAAACGAA 13912
 QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGTGAATGAAGCAGCATATCCGCTTGCAAGTTTG 13948
 RESULT 15
 D90735 18882 bp DNA linear BCT 07-FEB-1999
 LOCUS Escherichia coli genomic DNA. (22.3 - 22.7 mln).
 DEFINITION D90735 AB001340
 ACCESSION D90735.1 GI:4062539
 VERSION Complete and shotgun sequencing; hyaB; hyaC; hyaD; hyaE; hyaF;
 keywords appC; cyxA; cbaD; appB; cyxB; cbaB; appA; amsI; insA; insB; cspF;
 cspB; sfa; yccJ; yccM.
 SOURCE Escherichia coli (strain:K12) DNA, clone:Kohara clone #25.
 ORGANISM Escherichia coli
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE
 AUTHORS 1 (sites)
 Oshima, T., Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,
 Ikemoto, K., Inada, T., Itoh, T., Kajihara, M., Kanai, K., Kashimoto, K.,
 Kimura, S., Kitagawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K.,
 Mizobuchi, K., Mori, H., Motomura, K., Nakamura, I., Nishio, H.,
 Nishio, Y., Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K.,
 Wada, C., Yamamoto, Y., Yano, M. and Horiuchi, T.
 A 718-kb DNA sequence of the Escherichia coli K-12 genome
 corresponding to the 12.7-28.0 mln region on the linkage map
 JOURNAL DNA Res. 3 (3), 137-155 (1996)
 MEDLINE 97061202
 REFERENCE
 AUTHORS 2 (sites)
 Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiuchi, T.,
 Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K.,
 Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M.,
 Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H.,
 Motomura, K., Nakamura, I., Nishio, H., Nishio, Y., Oshima, T.,
 Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C.,
 Yamamoto, Y. and Yano, M.
 The systematic sequencing of the Escherichia coli genome in Japan

JOURNAL
REFERENCES
AUTHORS
TITLE
JOURNAL

Unpublished
3 (bases 1 to 18882)
Mori, H.
Submitted
Submitted (29-JUL-1996) Hirotsada Mori, NARA Institute of Science and Technology, Res. & Edu. Center for Genetic Info.: 8916-5 Takayama, Ikoma, Nara 630-01, Japan
(E-mail:hmori@gtc.aist-nara.ac.jp, Tel:81-7437-2-5660, Fax:81-7437-2-5669)
On Dec 24, 1998 this sequence version replaced gi:1651473.
Collaboration Information:
Project:
The Japan E.coli genome DNA sequencing project

COMMENT,
The Japan E.coli genome DNA sequencing project

Group:
The Japan E.coli genome DNA sequencing group

Members: (1995.4 - 1996.3)
Alba, H., Baba, T., Fujita, K., Hayashi, K., Honjo, A., Horiiuchi, T., Ikemoto, K., Inada, T., Isono, K., Isono, S., Itoh, T., Kanai, K., Kasai, H., Kashimoto, K., Kim, S., Kimura, S., Kitagawa, M., Kitakawa, M., Makino, K., Masuda, S., Miki, T., Mizobuchi, K., Mori, H., Motomura, K., Nakamura, Y., Nishimoto, H., Nishio, Y., Oshima, T., Saito, N., Sempel, G., Seki, Y., Tagami, H., Takemoto, K., Wada, C., Yamamoto, Y. and Yano, M.

Headed by:

Name: Takashi Horiiuchi
Address: National Institute of Basic Biology, Okazaki, 444, Japan
E-mail: kishori@nibb.ac.jp
Information operator:
Name: Hirotsada Mori

Address: NARA Institute of Science and Technology,

Ikoma, 630-01, Japan

E-mail: hmori@gtc.aist-nara.ac.jp

URL:

The Japan E. coli genome database

http:bsw3.aist-nara.ac.jp.

Location/Qualifiers

1. 18882

/organism="Escherichia coli"

/strain="K12"

/db_xref="taxon:562"

/map="22.3-22.7 min"

/clone="Kohara clone #225"

/note="Nucleotide position 1034630-1053511 from the

initiation site of *thra* (0 min.). This clone is from

Kohara Lambda miniset library."

1. 966

/gene="hyab"

<1. 966

/gene="hyab"

/note="ORF_ID:0224#6

similar to SwissProt Accession Number P19927"

/codon_start=1

/transl_table=1

/product="Hydrogenase-1 large chain (EC 1.18.99.1) (nife

hydrogenase) (membrane-bound hydrogenase 1 large subunit)

(hyal) "

/protein_id="BAA35738.1"

/db_xref="GI:4062540"

/transl_table=1

/note="ORF_ID:0225#1

similar to PIR Accession Number JY0074"

/codon_start=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

/transl_table=1

gene

CDS

NREBHVLAQAFRLGEGDLYGMLISRYAPDNHNVYAQYOAMRGALPOVAPEWSEFRI
 MVGCSLLIPLVLMIALVOITLRKIDQIRKIMKALMLPLPIALIEAGMPTERROP
 MAIODILPTYSANSAALTTGGLARSLIMIVGLYTLFLAEYLMQKVARLBSANQSRQ
 PTQOQG"
 5215..6351
 /gene="appb"
 /note="cbbd; cyxb"
 5215..6351
 /gene="appb"
 /note="ORF_id:0225#6
 similar to SwissProt Accession Number P26458"
 /codon_start=1
 /transl_table=1
 /product="Cytochrome bd-11 oxidase subunit 11 (EC
 1.10.3.-)"
 /protein_id="BAA35744.1"
 /db_xref="GI:1651480"
 /translation="MEDYETLRPIWMLLEIVLVFEMI SDGFDMGICLLPLVARND
 ERRIVINSVGAHMQVWLILAGALFAAMPRYVAAAFSGFYVAMLVLSLEFRL
 AFYRGRKISADARMRKMDAGLVIGSLVPVFGAFANLILGVEFAFATPOLRVEYLS
 FMOLLPEPLICGLSLGVNVLGGVWLOKTVGVHILRSOLATKRAALLVMCEFLA
 GYMLWGLDGFVLIAODANGSPNIMKIVLVIPGAMNNNTVESPVILFPIIGFPCRL
 LTVMAITRGRPMGLMASLMQREVITFAGITLFPFVMPSSVSPISSLTMDSTSSOL

Alignment Scores:

Pred. No.: 1.16e-162 Length: 18882
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 1 Gaps: 0

US-09-777-566a-2 (1-440) x D90735 (1-18882)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 |||
 DB 6536 ATGAACCGATCTTAATCCATTTTATCTCTCTGATTCGTTAACCCCGCAATCTGCA 6595
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValValIleValSerAlaHisGly 40
 |||
 DB 6596 TTCCCTAGAGTGGCCGAGCTGAAGCTGGAAGCTGGTGGATTTGTCAGTGCATAGGT 6655
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 |||
 DB 6656 GTGGGTGCTCCAAACCAAGCCACCACTGATGAGATGTCACCCAGACGATGGCCA 6715
 QY 61 ThrThrProValLysLeuGlyThrLeuThrProArgGlyGlyGluLeuIleAlaIleu 80
 |||
 DB 6716 ACCTGGCCGGTAAACCTGGTGGCTGACACCGCGGTGGTGAAGCTAATCGCTATCTC 6775
 QY 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 |||
 DB 6776 GGACATTAACCAACCCAGCTCTGTAGCCAGGATTTGCTGGGAAAAGGCTGCCCG 6835
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlyGlu 120
 |||
 DB 6836 CAGTCTGTGTCAGTGGCATATTGTCATGTCACAGCGGTACCCGTAANAACAGCGCAA 6895
 QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 |||
 DB 6896 GCCTTGCCCGCGGGCTGGCCACTGTGCAATACCGTACATACCCAGCAGATPACG 6955
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValLysGlnLeuAspAsnAla 160
 |||
 DB 6956 TCCAGTCCCGATCCGTTATTTAATCTCTTAANAACGTGGCTTGCCAACTGGATPACCG 7015
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 |||
 DB 7016 AACGTGACTGACGGGATCTTCACAGCAGGAGGATTAATTGCTGACTTACCGGGCAT 7075
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
 |||
 DB 7076 CGGCAAAAGCGGCTTCCGAACTGGAACGGGCTCTAATTTCCGCAATCAAACTTGTC 7135

QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 |||
 DB 7136 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGATTAACCATCGGAACCTC 7195
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 |||
 DB 7196 AAGGTGAGCGCGGCACATATGCTCATTAACCGGTGGGTAAAGCCTCGCATCAATGCTGACG 7255
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTTPGlyArgIleThr 260
 |||
 DB 7256 GAGATATTTCTCCGCAACAAGCAGAGGATGCGGAGCGGGGTGGGAGAGATCAC 7315
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 |||
 DB 7316 GATTCAACACCACTGGAACACCTCTCTAAGTTTGATTAACGCGCAATTTTATTCCTCAAA 7375
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 |||
 DB 7376 CGCAGCGCAGAGGTGGCCCGCAGCGCGCCACCCGTTATTAGATTGATCAAGACAGCG 7435
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 |||
 DB 7436 TTGACGCCCATCCACCGCAAAAACAGCGGTATGTGTGACATTACCCTTCAGTCTG 7495
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 |||
 DB 7496 TTTATCCCGCGACACGATACTAATCTGCAAAATCTCGCGCGCATGAGACTCAACTGG 7555
 QY 341 ThrLeuProGlnGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 |||
 DB 7556 ACGCTCCCGGTACGCGGATTAACACCGCCGACAGTGGTGAACGTGTTGAACGCTGG 7615
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 |||
 DB 7616 CGTGGCTTAACCGATTAACAGCAGCTGATTCAGTTTCGCTTCACAGACTTTACAG 7675
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 |||
 DB 7676 CAGATGGCTGATTAANAACCGCTCTCATTAATACGCCGCCGGAAGAGTGAACCTGAC 7735
 QY 401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 |||
 DB 7736 CTGGCAGAGATGTGAAGAGCGAAATGCCAGGCGCATGTTCGTTGCGCAGGTTTACGCAA 7795
 QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
 |||
 DB 7796 ATCGTGAATGAAGCACCACATACCGCGGTGCAGTTTG 7831

Search completed: June 12, 2003, 14:59:54

Job time : 3669 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 12:09:10 ; Search time 328 Seconds

(without alignments)
3020,973 Million cell updates/sec

Title: US-09-777-566a-2

Perfect score: 2302
Sequence: 1 MKAIIPLFLSLIPLFPOSA.....IYNFARIPACSLRSHHHHH 440

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-O/cgnt2.1/USPRO.spool/US09777566/runat_05062003_154807_6525/app.query.fasta.1.583
-DB=N_Geneseq.101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPEXT=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi
-LIST=45 -DOCAIIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09777566 @CGN.1.1.263 @runat_05062003_154807_6525 -NCPU=6 -ICPU=3
-NO.MMAP -LARGEOBFRY -NR3_SCORES=0 -NAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -DelEXT=7

Database :

N_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	22 AAC88885	Escherichia coli B
2	2302	100.0	1323	24 AAC86473	Escherichia coli B
3	2302	100.0	1323	24 AAD25460	Escherichia coli B
4	2235	97.1	1299	21 ABK12514	DNA encoding phyt
5	2235	97.1	1901	24 AAD25463	Escherichia coli a
6	2235	97.1	3470	21 AAC68298	R15/APPA plasmid c
7	2235	97.1	4060	21 AAC68296	R15/APPA plasmid c
8	2235	97.1	6116	21 AAC68297	R15/APPA plasmid c
9	2235	97.1	6708	21 AAC68295	R15/APPA plasmid c
10	2235	97.1	17732	21 AAC68300	LamA2/APPA plasmid
11	2235	97.1	20623	21 AAC68294	LamA2/APPA plasmid
12	2221	96.5	1489	21 AAA28216	E. coli acidic pho
13	2217	96.3	1486	22 AAD06831	E. coli acidic pho
14	2205	95.8	5421	21 AAC68299	SV40/APPA plasmid
15	2197	95.4	1486	22 AAD06832	E. coli acid phosph
16	2194.5	95.3	1272	20 AAX26340	DNA encoding an Es
17	1401.5	60.9	846	22 AAS41940	Genomic sequence #
18	1196.5	52.0	1737	23 AAS89885	DNA encoding novel
19	813	35.3	519	22 AAS41139	CDNA encoding novel
20	759.5	33.0	4050	23 AAS64582	DNA encoding novel
21	657	28.5	1264	23 AAS88443	DNA encoding novel
22	657	28.5	1264	23 AAS94316	DNA encoding novel
23	539	23.4	11710	24 AAB43718	DNA encoding novel
24	378.5	16.4	356	23 AAS89883	E. coli genome sec
25	352	15.3	5975	15 AAO55758	Klebsiella pneumon
26	274	11.9	5976	23 AAS88450	DNA encoding novel
27	229	9.9	898	23 AAS75019	DNA encoding novel
28	195.5	8.5	1282	23 AAS87535	DNA encoding novel
29	195	8.5	1281	24 AAB17591	DNA encoding novel
30	195	8.5	1352	24 AAB14473	Human protein phos
31	184.5	8.0	3189	23 AAS82488	DNA encoding novel
32	184.5	8.0	3189	23 AAS89136	DNA encoding novel
33	166	7.2	427	21 AAA67371	Pinus radiata gluc
34	136.5	5.9	1272	24 ABL36298	Human lysosomal ac
35	133	5.8	1584	23 ABL11483	Drosophila melanog
36	124.5	5.4	1197	23 ABL29887	Drosophila melanog
37	119	5.2	1398	22 AAD17943	Aspergillus fumiga
38	119	5.2	1410	23 ABL09027	Drosophila melanog
39	119	5.2	1571	19 AAX03144	Aspergillus fumiga
40	119	5.2	1571	20 AAZ27422	A. fumigatus phyt
41	117	5.1	1588	18 AAT72721	PAP-GM-CSF immuno
42	117	5.1	3061	19 AAV29653	Human prostate can
43	117	5.1	3061	20 AAX26062	US5882864 Seq ID N
44	117	5.1	3061	21 AAZ87547	Prostatic acid pho
45	117	5.1	3061	22 AAS03766	Biomarker UC band

ALIGNMENTS

RESULT 1

AAC88885 standard; DNA; 1323 BP.

AAC88885;

07-MAR-2001 (first entry)

Escherichia coli B phytase enzyme nucleotide sequence.

Escherichia coli B; phytase enzyme; anabolic; phytate digestion;

nutrition; ds.

Escherichia coli.

30-NOV-2000.

XX 15-MAY-2001; 2001WO-US15764.
 PF
 XX 25-MAY-2000; 2000US-0580937.
 PR
 XX (DIVE-) DIVERSA CORP.
 PA
 XX Short JM, Kretz KA, O'Donoghue E;
 PI
 XX MPI: 2002-164149/21.
 DR P-PSDB; AAE22836.
 DR
 XX New dietary aids comprising sustained release biocompatible
 PT compositions, comprise agent that assists in digestion, useful for
 PT delivering enzymes, therapeutics, medicine or agents to an organism
 XX
 PS Disclosure; Fig 1; 89pp; English.
 XX
 CC The present invention relates to novel dietary aids comprising sustained
 CC release biocompatible composition which comprises an agent (enzymes such
 CC as phylase, amylase, esterase, protease) that assists in digestion. The
 CC biocompatible composition is effective upon oral consumption and release
 CC in the digestive tract of a subject. The dietary aids are useful for
 CC delivering enzymes, therapeutics, medicine and agents to an organism.
 CC The use of enzymes and other agents in digestive aids of livestock or
 CC domesticated animals not only improves the animal's health and life
 CC expectancy but also assists in increasing the health of livestock or
 CC in the production of foodstuffs from livestock. The present sequence
 CC is Escherichia coli phylase DNA.
 XX
 SQ Sequence 1323 BP; 323 A; 353 C; 357 G; 289 T; 1 other;
 S0
 Alignment Scores:
 Pred. No.: 5.79e-202 Length: 1323
 Score: 2302.00 Matches: 440
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 24 Gaps: 0
 US-09-777-566a-2 (1-440) x AAD36473 (1-1323)
 QY 1 MetLysAlaIleuLeuSerPheLeuSerLeuLeuLeuProLeuThrProGlnSerAla 20
 DB 1 ATGAAGGATCTTAATCCATTTTATCTTCGATTCGCTAACCCCAATCTCA 60
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 61 TTCCCTCAGAGTGAAGCCGAGCTGAAGCTGGAAGTGTGATGTCAGTCGATGCT 120
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 121 GTGGGTGCTCCACCAAGCCAGCCGACGACTGATGCGATGTCACCCAGAGCATGGCCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
 DB 181 ACGGCGCGGTAAGTAACGTGTGGCTGACACCGCGNGGTGGTGAAGCTAATGCCATCTC 240
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 DB 241 GGACATATACCAAGCCAGCGCTGTGTAGCCAGGATGTGGCGAAGGGGTGCCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluValGlyThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGGTGCGCATTAATGCTGATGTGCGAGCGCTACCCGTAACAGCGCAA 360
 QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTCGGCGCGCGGTGGACATGCTGACATGCAATACCGTACATACCGACATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGGATCGGTATTTAATCCCTTAATAAACTGCGGTTGGCAACTGGAATACCGC 480

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTGACTGACCGGATCTTACGACGAGGAGGATCAATGCTGACTTACCGGGCAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnProGlnSerAsnLeuCys 200
 DB 541 CGGCAAAAGCGGCTTCCGGAAGCTGAAGCGGGTGTCTAATTTCCGCAATCAACTGTGCG 600
 QY 201 LeuLysArgGlyLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
 DB 601 CTTAAGCTGGAAGACAGAGCAAGAGCTGTTCAATTAAAGGAGCATTAACCAATCGAATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTGACCGCGGACATATGCTCTAATACCGGTGCGGTAACCTTCGATCAATCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr 260
 DB 721 GAGATATTTCTCTGCAACACAGACAGGAAATGCCGCGGGGTGGGAAGGATCAC 780
 QY 261 AspSerHisGlnTyrPheThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuGln 280
 DB 781 GATTCACACACAGTGAACACCTTCTAAGTTGATTAACGCGCAATTTATTTCTCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCGACAGGTGGCCGACCGCGCCGACCCGCTTATGATTCATCATGACGACG 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATTCACACCCCAAAACAGCGCTATGCTGTGACATTAACCATTCAGTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnThr 340
 DB 961 TTTATTTGCCGACACGATACTAATCTGCAATCTCGGCGCGCAGCTGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluThr 360
 DB 1021 ACGCTCCCGGTACCGCGGATTAACAGCGCGGAGTGTGTAAGTGTGTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTAAGCATTAACAGCAGCAGATGAGTTCGCTGCTCCACATCTTACAG 1140
 QY 381 GlnMetLysAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
 DB 1141 CAGATGCGGTATTAACCCGCTGCTAATTAATACGCGCGCGGAGAGTGAACCTGACC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGATGTGAAGACGCAAAATGCGCAGGCGATGTGTTGGCGGTTTACGCA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
 DB 1261 ATCGGATGAAGCAGCATACCGCGCTGAGTTGAGATGTCATCACACATCACATCAC 1320
 RESULT 3
 AAD25460
 ID AAD25460 standard; DNA; 1323 BP.
 AC
 AAD25460;
 XX 26-MAR-2002 (first entry)
 XX
 DE Escherichia coli B phylase DNA.
 XX
 KW Bacterial phylase: K12 appa phylase; protease stability; anabolic;
 KW gastrointestinal; nutritional value; feed treatment process; therapy;
 KW thermal tolerance; growth performance; alcoholic drink; diopulping;
 KW non-alcoholic drink; biobleaching; B phylase; ds.
 XX
 OS Escherichia coli.

RESULT 4

ABK12514 standard; DNA; 1299 BP.

AC ABK12514;

05-JUN-2002 (first entry)

DNA encoding phytase associated protein.

Phytase; gene: ds.

Unidentified.

FT	Key	Location/Qualifiers
FT	CDS	1..1299
FT	/*tag=	a
FT	/product=	"Phytase associated protein"

KR9086028-A.

15-DEC-1999.

25-MAY-1998; 98KR-0018810.

25-MAY-1998; 98KR-0018810.

(WO0J-) WO0JIN CO LTD.

Bae HD, Forceburgh CW, Goloben S, Cheng KJ;

WPI: 2000-645078/62.

P-PSDB; AAU77775.

Novel phytase gene, recombinant phytase and usage thereof -
 Claim 1; Fig 2; 10pp: Korean.

The invention relates to a novel phytase gene, a recombinant phytase gene and their uses. This sequence encodes a phytase associated protein, described in the invention.

SQ Sequence 1299 BP; 317 A; 345 C; 355 G; 282 T; 0 other;

Alignment Scores:

Pred. No.:	8,14e-196	Length:	1299
Score:	2235.00	Matches:	430
Percent Similarity:	99.54%	Conservative:	0
Best Local Similarity:	99.54%	Mismatches:	2
Query Match:	97.09%	Indels:	0
DB:	21	Gaps:	0

US-09-777-566a-2 (1-440) x ABK12514 (1-1299)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
 DB 1 ATGAAGGAGATCTTAATCCCATTTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 60
 QY 21 PheAlaGlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgIleGly 40
 DB 61 TTCGCTCAGAGTACCGGAGCTGAAGCTGGAATGTGCTGATGTCTCATGTATGTT 120
 QY 41 ValAlaGlnAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 121 GTGCTGCTCCAAACCAAGGCCACGCACTGATGTCAGATGTCCACCCAGACGATGGCA 180
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80
 DB 181 ACCTGGCCGTAACCTGGGTGGCTGACACCGCGGTGGTGAAGCTAATCGCTATCTC 240
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysGlyPro 100

DB 241 GGACATTAACCAAGCCAGCGTCTGGTACCGCAGGATGTCGGCAAAAAGGCTGCCG 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 301 CAGTCTGGTCAGGTGCGATTAATTCGTATGTCAGACGACGATCCCGTAACAGGCGAA 360
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 361 GCCTTGGCGCGCGGCTGGCACTGACGTGCAATTAACGATACCAAGCAGGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 421 TCCAGTCCCGATCCCTTAATTAATCTTAATAAAGCTGGCTTGGCAACTGGATTAACCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 481 AACGTACTGACGCGATCCTCAGCAGGCGAGGAGGTCATTAATGCTGACCTTAACCGGAT 540
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 541 CGGCAAGCGCGTTCGGGAACCTGGAACGGGTGCTTAATTTCCGCAATCAAACTTGTC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 601 CTTAAACGTGAGAAACAGGACGAAAGCTGTTCAATTAACGACGATTAACCAATCGGAATC 660
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyValValSerLeuAlaSerMetLeuThr 240
 DB 661 AAGGTAGCGCGCGCAATGTCATTAACCGGTGGGTAGCGTTCGCAATGCTGACG 720
 QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyValArgIleThr 260
 DB 721 CAGATATTTCTCTCCCAACCAACAGCAGGAAACCGGAGCCGGGTGGGGAAGATCAAC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
 DB 781 GATTCACACAGTGGAGAACCTTGTGTAATTTGCAATACGGCAATTTTAATTTGCTACAA 840
 QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 841 CGCAGCCAGAGGTGGCCGCGCAGCGCCGCGCACTTATTAATTAATTAATTAATTAATTA 360
 QY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
 DB 901 TTGACGCCCATCCACCGCAAAACAGGCGTATGCTGACATTAACCACTCACTGCTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
 DB 961 TTTATCGCGCGACAGATACTATCTGCAATCTCGCGCGGCACTGAGCTCAACTG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGluArgTrp 360
 DB 1021 ACGTTCGCCGTCACCGGATTAACAGCGCGCGGATGATGTAAGCTGTTTGAACGCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 1081 CGTCGGCTTAAGGATTAACAGCCAGATGATCAAGTTCAGTTCCTTCCAGACTTAACG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
 DB 1141 CAGATGCGGTAATAAAGCGCGCTGCAATTAATTAACGCGCGGAGAGTGAAGTGAAC 1200
 QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 1201 CTGGCAGGATGTGAAGAGCGAAATGCGAGGCGATGTGCTGGCAGGTTTATACGAA 1260
 QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 1261 ATCGTAATGAAGCACCGCATACCGGCGTGCAGTTTG 1296

RESULT 5

AAD25463 standard; DNA; 1901 BP.

XX AAD25463;
XX 26-MAR-2002 (first entry)
XX Escherichia coli appa phytase wild type DNA.
XX
XX Bacterial phytase; K12 appa phytase; protease stability; anabolic;
XX gastrointestinal; nutritional value; feed treatment process; therapy;
XX thermal tolerance; growth performance; alcoholic drink; biopulping;
XX non-alcoholic drink; biobleaching; ds.
XX
XX Escherichia coli.
XX
XX Key Location/Qualifiers
XX CDS 188..1486
XX FT /*tag= a
XX FT /product= "E. coli appa phytase protein"
XX
XX MO200190333-A2.
XX
XX 29-NOV-2001.
XX
XX 24-MAY-2001; 2001MO-US17118.
XX
XX 25-MAY-2000; 2000US-0580515.
XX
XX (DIVE-) DIVERSA CORP.
XX
XX Short JM, Kretz KA, Gray KA, Barton NR, Garrett JB, O'Donoghue E;
XX
XX WPI: 2002-083108/11.
XX
XX P-PSDB: AAE15807.
XX
XX New bacterial phytase for e.g. improving the nutritional value of
XX phytate-containing foodstuffs and subsequently improving the growth
XX performance of an organism that consumes it, or in treating animal
XX digestive systems -
XX
XX Claim 42; Fig 7; 170pp; English.
XX
XX The patent discloses recombinant bacterial phytase from Escherichia coli
XX K12 appa phytase. The enzyme has phytase activity and improved thermal
XX tolerance when compared with wild-type phytase. It has improved protease
XX stability at low pH. The recombinant phytase is useful for improving the
XX nutritional value of phytate-containing foodstuffs and subsequently
XX improving the growth performance of an organism that consumes it, in
XX treating animal digestive systems, in feed treatment processes and for
XX in vitro purposes related to research, discovery and development. They
XX are also used for generating recombinant digestive system life forms,
XX for producing or manufacturing alcoholic and non-alcoholic drinks based
XX on the use of moulds, grains and/or plants, in biopulping and bio-
XX bleaching where a reduction in the use of environmentally harmful
XX chemicals that are traditionally used in the pulp and paper industry
XX is desired and in the reduction of possible elimination of the need
XX for mineral supplements, enzymes or therapeutic drugs for animals
XX from the daily feed thus increasing the amount calories and nutrients
XX present in the feed. The present sequence is a DNA encoding E. coli
XX appa phytase wild type protein.
XX
XX Sequence 1901 BP: 474 A; 499 C; 499 G; 428 T; 1 other:
XX
XX Alignment Scores:
XX Pred. No.: 1.38e-195 Length: 1901
XX Score: 2235.00 Matches: 430
XX Percent Similarity: 99.54% Conservative: 0
XX Best Local Similarity: 99.54% Mismatches: 2
XX Query Match: 97.09% Indels: 0
XX DB: 24 Gaps: 0
XX
XX US-09-777-566a-2 (1-440) x AAD25463 (1-1901)
XX
XX 1 MetysAlaIleuIleProPheIeuSerIeuIleProIleuThrProGlnSerAla 20

DB 188 ATGAAGGCGATCTTATCCCATTTTATCTCTTCGATTCCTGTAACCCGCAATCTCA 247
QY 21 PheAlaGlnSerGluProGluIleuLysLeuGluSerValIleValSerArgHisGly 40
DB 248 TTCGCTCAGAGTACCCGAGCTGAAAGCTGGAAATGTGTCATGTCATGTCATGTCAT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
DB 308 GTGCGTGTCCCAACCAAGGCCAGCACTGATGATGATGATGATGATGATGATGATGAT 367
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlyGlyGlyGlyGlyGly 80
DB 368 ACCGTGCGGTAACCTGGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 427
QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB 428 GGACATTCACCAAGCCGACGCTGTGATGACCGAGATTCGTGCGCAAAAAGGCGTGGCCG 487
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
DB 488 CAGTCTGTGAGTGGCGCATTTATGCTGATGTCGACGAGCGATCCGTAAACAGCGCA 547
QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB 548 GCCTTGCGCGCGGCTGGCTGACCTGCTGCAATACCGTACATACCGAGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB 608 TCCAGTCCCGAGATCCGTTATTTATTCCTCAAAAACATGGCGTTGGCAATGATACGCG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
DB 668 AACGTGACTGACGCGCATCTCTCAGCAGCGCAGAGGCTCAATTGCTGACTTTACCGCGCAT 727
QY 181 ArgGlnThrAlaPheArgGluLeuGluIleuArgValLeuAsnPheProGlnSerAsnLeuGly 200
DB 728 CGGCACAGCGGCTTTCCGCACTGCAACGCGGCTGCTTAATTTCCCAATCAAACTGTGTC 787
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB 788 CTTAAACGTGAGAAACAGAGCAAAAGCTGTCATTACGACGAGCATTCATCCGAGACTC 847
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB 848 AAGGTGACGCGCGCAATGTCATTAAACCGTGGGTAAGCTTCGATCATGCTGACG 907
QY 241 GluIlePheLeuLeuGlnAlaGlnIleuMetProGluProGlyTrpGlyArgGlyLeuThr 260
DB 908 GAGATATTTCTCTCAACACAGCACAGGAAATGCGGAGCGGGGTGGGAAGGATCACAC 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
DB 968 GATTCACACCAAGTGAACACCTGTGTAAGTTGCAATACGCGCATTTTATTTGCTACAA 1027
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB 1028 CGCAGCCAGAGGTGCGCGCAGCGCCGCAACCCGCTTATAGATTATGATCAAGACACG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaIleArgValThrLeuProThrSerValGlu 320
DB 1088 TTGACGCCCATCCACCGCAAAAACAGCGATGATGATGATGATGATGATGATGATGATG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuValAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
DB 1148 TTTATACGCCGAGACAGATACATATGTCGCAATATCGCGCGCAGCTGAGCTCAACATCG 1207
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlyGlyGlyGlyGlyGly 360
DB 1208 ACGCTCCCGGTGACCGCGATTAACACGCGCGCAGGTGATGAATGCTGTGTAACGCGTGG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

Db	1268	CGTCGGCTAAGCATATACAGCCAGTGCATTCAGGTTTGGCTGGTCTTCACAGCTTTACAG	1327
Qy	381	GlMetCArgAspIysThrProLeuSerLeuAsnThrProTroglyValIysLeuThr	400
Db	1328	CAGATGGCATATAAAGCGCGCTGTCATTAAATACGCCCGGAGAGGTGAACGTACC	1387
Qy	401	LeuAlaGlyCysGluIuArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln	420
Db	1388	CTGGCAGGATGTAAAGAGCAAAATGCGCAGGCGATGTGTCGTGGCAGGTTTACGCA	1447
Qy	421	IleValAsnGluAlaArgIleProAlaCysSerLeu	432
Db	1448	ATCGTGAATGACACGACGATACCGGCTGCAGTTTG	1483
RESULT 6			
AAAC68298			
AC	AAAC68298	standard; DNA; 3470 BP.	
XX	AAAC68298;		
XX	20-FEB-2001	(first entry)	
XX			
XX	RI5/APPA	plasmid coding sequence.	
DE			
KM	Transgenic animal; salivary protein; phytase; phosphorus; animal growth;		
KM	environmental pollution; pig; ds.		
XX	Rattus sp. - chimeric.		
OS	Escherichia coli - chimeric.		
XX	MO200064247-A1.		
PN			
XX	02-NOV-2000.		
PD			
XX	20-APR-2000; 2000MO-CA00430.		
PF			
XX	23-APR-1999; 99US-0130508.		
PR			
XX	(UYGU-) UNIV GUELP.		
PA			
PI	Forsberg CW, Golovan S, Phillips JP;		
XX			
DR	WPI: 2000-687245/67.		
XX	P-PSDB; AAB36261.		
PT	Transgenic non-human animal for gastrointestinal tract specific		
PT	expression of a protein, preferably phytase, comprises a nucleic acid		
PT	sequence including a heterologous transgene construct encoding the		
PT	protein		
XX			
PS	Claim 14; Fig 21; 152pp; English.		
XX			
CC	The present invention provides transgenic animals which produce desired		
CC	proteins, in this case pigs which expresses phytase in the salivary		
CC	gland. Low phytase production levels result in phytate in the diet being		
CC	excreted and causing phosphorus contamination in water, as well as		
CC	reducing the growth of animals. The invention provides a number of		
CC	transgenes containing the E. coli APPA phytase coding sequence.		
XX			
SO	Sequence 3470 BP; 1065 A; 721 C; 735 G; 949 T; 0 other:		
Alignment Scores:			
Pred. No.:	3,16e-195	Length: 3470	
Score:	2235.00	Matches: 430	
Percent Similarity:	99.54%	Conservative: 0	
Best local Similarity:	97.09%	Mismatches: 2	
Query Match:	21	Indels: 0	
DB:		Gaps: 0	
US-09-777-566A-2 (1-440) x AAC68298 (1-3470)			
Qy	1	MethylalalleuIleuIleuProPheLeuSerLeuLeuIleuProLeuThrProGlnSerAla	20
Db			
Db	1811	ATGAAGGCATCTTAATCCCATTTTATATCTTCATTCGATTACCGGCAATCTGCA	1871
Qy	21	PheAlaGlnSerGluProGluLeuIysLeuGlnSerValIleValSerArgHisGly	40
Db	1871	TTTCGTCAGAGTATGACGGAGGTGAAGCTGGAAGATGTGGTATGTTCAGTCGCAATGT	1930
Qy	41	ValArgAlaProThrIysAlaThrGlnLeuMetGlnAspValIleThrProAspAlaThrPro	60
Db	1931	GTCCGCTCTCAACCAAGGCCACGAACTGAAGAGATGTACCCCGACGAGCATGGCCA	1990
Qy	61	ThrTrpProValIysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaIleu	80
Db	1991	ACCTGGCCGGTAAATATGGGTGGCTGACACCCGCCGGTGGAGACTAATGCCCTATCTC	2050
Qy	81	GlyHisIleGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIysGlyCysPro	100
Db	2051	GGACATTACCAACGCCAGCGTCTGGTAGCCAGCATTTGCTGGCAAAAGCGTCCCG	2110
Qy	101	GlnSerGlyGlnValAlaIleIleAlaAspValAspGluValArgIleThrGlyGlu	120
Db	2111	CAGTCTGGTCAGGCTCCGATTTATGTGTATGTGCAGAGCGGTACCCGTAACAGCGCA	2170
Qy	121	AlaPheAlaIaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	2171	GCCTTGCCCGCCGGCTGGCACCTGTACTGTGCATAATACGTCATACCAAGCAGATACG	2230
Qy	141	SerSerProAspProLeuPheAsnProLeuIleThrGlyValCysGlnLeuAspAsnAla	160
Db	2231	TTCAGTCCGATCCGATTATTATTCCTTAATAACTGGGTTGGCAACTGGATACGCG	2290
Qy	161	AsnValIleThrAspAlaIleLeuSerArgAlaGlyIleSerIleAlaAspPheThrGlyHis	180
Db	2291	AACGTGACTACCCGATCTCAAGCAGGCGGAGGTCATATGTGACTTACCGGGCAT	2350
Qy	181	ArgGlnThrAlaPheArgIleuGlnIuArgValIleuAsnPheProGlnSerAsnLeuIys	200
Db	2351	CGGCAAAAGCGGTTTGGCAACTGGAACGGGTCTTAATTTCCGCAATCAAACTGTGGC	2410
Qy	201	LeuIysArgGluIysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu	220
Db	2411	CTTAACGTCAGAAACAGGACGAACCTTTCATTATTAAGCAGCATTTACATCGGAATCTC	2470
Qy	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	2471	AAGGTGAGCCCGACATGTCTCATTAACCGGTGGCGTAACTGCATCATCTCTGACG	2530
Qy	241	GluIlePheLeuLeuGlnAlaGlnGlyMetProGluIuProGlyIleArgIleThr	260
Db	2531	GAGATATTTCTCTGCAACACAGCACGGAAGCCGCGGTGGGAGAGATCAACC	2590
Qy	261	AspSerIleGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyLeuLeuGln	280
Db	2591	GATTACACACAGGAAACACCTTTCATAGTTTGCATACGCGCAATTTATTTCGTCACA	2650
Qy	281	ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	2651	CGACCGCAAGAGTCCCGCAGCCGCCGACCCGCTTATATGAT	


```

Oy 381 GlnmetAgaApLyThProLeuSerLeuAnThrProPGlyGValLySerThr 400
Db 2951 CAGATGCGTGTATMAAACCCCGCTGTCATTAAATACGCCGCCGGAGAGTGAACTGACC 3010
Oy 401 LeuAlaGlyCysGluGluAraGAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 3011 CTGGCAGATGTGAAGACGCGAAATCGCGAGGCGCAGTCGTGCTTGCGAGGTTTACGCCAA 3070
Oy 421 TLeuAlaGlnGluAlaArgTlleProAlaCysSerLeu 432
Db 3071 ATCGTAATGAAGCAGCATACCGCGCTTGAGATTGG 3106

RESULT 7
AAC68296
ID AAC68296 standard; DNA; 4060 BP.
XX
AC AAC68296;
XX
DT 20-FEB-2001 (first entry)
XX
DE R15/APPA plasmid coding sequence.
XX
KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX environmental pollution; pig; ds.
XX
OS Rattus sp - chimeric.
XX
OS Escherichia coli - chimeric.
XX
PN WO200064247-A1.
XX
PD 02-NOV-2000.
XX
PF 20-APR-2000; 2000MO-CA00430.
XX
PR 23-APR-1999; 99US-0130508.
XX
PA (UYGU-) UNIV GUELPH.
XX
PI Forsberg CW, Golovan S, Phillips JP;
XX
DR WPI: 2000-687245/67.
XX P-PsDB; AAB36259.
XX
PT Transgenic non-human animal for gastrointestinal tract specific
PT expression of a protein, preferably phytase, comprises a nucleic acid
PT sequence including a heterologous transgene construct encoding the
PT protein.
XX
XX Claim 14; Fig 19; 152pp; English.
XX
PS
XX
CC The present invention provides transgenic animals which produce desired
CC proteins, in this case, pigs which expresses phytase in the salivary
CC gland. Low phytase production levels result in phytate in the diet being
CC excreted and causing phosphorus contamination in water, as well as
CC reducing the growth of animals. The invention provides a number of
CC transgenes containing the E. coli APPA phytase coding sequence.
XX
SQ Sequence 4060 BP; 1257 A; 814 C; 843 G; 1146 T; 0 other.

Alignment Scores:
Pred. No.: 3,93e-195 Length: 4060
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 99.54% Mismatches: 2
Query Match: 97.09% Indels: 0
DB: 21 Gaps: 0

US-09-777-566A-2 (1-440) x AAC68296 (1-4060)
Oy 1 MettysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1811 ATGAAGCCATCTTAAATCCATTTTAACTCTTCGTGATTCGGTTAAACCCGCAATCTGCA 1870

```

QY	21	PheAlaGlnSerGlnupProGlnLeuLeuLysLeuGlnLysSerValValIleValSerArgHisGly	40
Db	1871	TTGGCTCAGAGTGAAGCCGAGACTGAGCTGAGAAAGTGTGTGATTTGTCAGTCTCATGCT	1930
QY	41	ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValIThrProAspAlaTrpPro	60
Db	1931	GTGGCGTCTCCAAACCAAGCCACGCAACTGATGACAGATGTGCACCCCGAGACGATGGCGCA	1990
QY	61	ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrLeu	80
Db	1991	ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGGGGTGGTGAAGCTATATGCGCTATCTC	2050
QY	81	GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro	100
Db	2051	GGACTTACCAACGCGCACGCTGTTGTTACCGACGAGATGCTGGCGAAAGAGGCTGCCCG	2110
QY	101	GlnSerGlyGlnValAlaIleIleLeuAlaAspValAspGlnArgThrArgLysThrGlyGln	120
Db	2111	CAGCTGTGTCAGGTGCGCATTAATGTGATGTGACGCGAGCTACCCCGTAAACAGGCGAA	2170
QY	121	AlaPheAlaIaAGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr	140
Db	2171	GCCTTCCGCCCGCGGCTGGCACCTGACTGTGCATTAACCGTACATACCAGCAGATACG	2230
QY	141	SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla	160
Db	2231	TTCAGTCCCGGATCCGTTATTTAACTCTTAAAACTGGCGCTTCCAACTGATACACGG	2290
QY	161	AsnValThrAspAlaIleLeuSerArgLysGlySerIleAlaAspPheThrGlyHis	180
Db	2291	AACGTGACTGACCGCATCTTACGACGAGCAGAGAGGTAAATTGCTGACTTATACGGGCAT	2350
QY	181	ArgGlnThrAlaPheArgGlnLeuGlnLuarGlyValLeuAsnPheProGlnSerAsnLeuGly	200
Db	2351	CGCGAAACGGCGTTTCGGCACTGGAAACGGGTGCTTATTTCCGCAATCAACTTGGTC	2410
QY	201	LeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaIleuProSerGlnLeu	220
Db	2411	CTTAAACCTGAGAAACAGAGCAAAAGCTGTTCAATTAAACGAGGATTAACATCTGGAACCTC	2470
QY	221	LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr	240
Db	2471	AAGGTAGACGCCGCAATATGCTCATTTAACCGGTGCGGTAAACCTCGCATCAATGCTGACG	2530
QY	241	GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnupProGlyTyrPglyArgIleThr	260
Db	2531	GAGATATTTCTCTGCAACAACAGACAGAGATGCCGCGAGCGGGGTGGGGAAGATCACCC	2590
QY	261	AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln	280
Db	2591	GATTCACACACAGTGGAAACCTCTGTAAAGTTGTGCAATACGGGCAAATTTATTCTTCACA	2650
QY	281	ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla	300
Db	2651	CGACAGCCAGAGGTTGGCCCGCAGCGCGCCACCCCGTTATTAAGATTGATCAAGACAGCG	2710
QY	301	LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu	320
Db	2711	TTTGACGCCCTCCACCGCAAAACAGCGCATGTGTGTGACATTAACCCACATTCAGTCTGTG	2770
QY	321	PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp	340
Db	2771	TTTATCCGCCGACACGATACTAATCTGGCAAAATCTCGCGCGCGACCTGGAGCTCAACTGG	2830
QY	341	ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnLuarGlyTrp	360
Db	2831	ACGCTTCGCCGTACCGGATTAACACGCCCGCCAGGTGTGAACTGATGAGTGTGAACGCTGG	2890
QY	361	ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln	380
Db	2891	CGTGGCGCTAACCGATTAACAGCCAGTGTGAATTCAGTTTCGCTGGTGTTCACAGACTTAAACG	2950


```

Db      2951 CAGATCGCTGATTAACGCCCTGTCATTAAATACGCCGCCGAGAGAGTGAACCTGACC 3010
OY      401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db      3011 CTGCAGAGATGTGAAGACGAAATGCGCAGGCGCATGTGTGCTGGCAGGTTTATTCGCA 3070
OY      421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db      3071 ATCGTGAATGAAGCAGCATACCGCTGTCAGATTGG 3106

RESULT 9
AAC68295
ID      AAC68295 standard; DNA; 6708 BP.
XX
XX      AAC68295;
AC      20-FEB-2001 (first entry)
DT
XX
XX      R15/APPA plasmid coding sequence.
DE
XX      Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
KW      environmental pollution; pig; ds.
XX
XX      Rattus sp - chimeric.
OS      Escherichia coli - chimeric.
XX
XX      WO200064247-A1.
PM
XX      02-NOV-2000.
PD
XX      20-APR-2000; 2000WO-CA00430.
PF
XX      23-APR-1999; 99US-0130508.
PR
XX      (DYGU-) UNIV GUELPH.
PA
XX      Forsberg CM, Golovan S, Phillips JP;
PI
XX      WPI: 2000-687245/67.
DR      P-PSDB; AAB36258.
XX
XX      Transgenic non-human animal for gastrointestinal tract specific
PT      expression of a protein, preferably phytase, comprises a nucleic acid
PR      sequence including a heterologous transgene construct encoding the
PT      protein.
XX
XX      Claim 56; Fig 18; 152pp; English.
PS
XX      The present invention provides transgenic animals which produce desired
CC      proteins, in this case pigs which expresses phytase in the salivary
CC      gland. Low phytase production levels result in phytate in the diet being
CC      excreted and causing phosphorus contamination in water, as well as
CC      reducing the growth of animals. The invention provides a number of
CC      transgenes containing the E. coli APPA phytase coding sequence.
XX
XX      Sequence 6708 BP; 1916 A; 1479 C; 1515 G; 1798 T; 0 other;

Alignment Scores:
Pred. No.: 7,86e-195 Length: 6708
Score: 2235.00 Matches: 430
Percent Similarity: 99.54% Conservative: 0
Best Local Similarity: 97.09% Mismatches: 2
Query Match: 21 Indels: 0
DB: 21 Gaps: 0

US-09-777-566a-2 (1-440) x AAC68295 (1-6708)
OY      1 MetysAlaIleIleuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db      1811 ATGAAAGCATCTTAATCCATTATTTATCTTCTGATTCCGTTAAACCCCGCAATCTGCA 1870
OY      21 PheAlaGlnSerGluProGluLeuLysGlnSerValIleValSerArgHisGly 40

```

```

Db      1871 TTCGTCAGAGTGAAGCCGAGACTGAGCTGGAAGTGAGTGATTTGTCAGTCGTCATGCT 1930
OY      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      1931 GTGGCTGCTCCAAACCAAGCCAGCCAGCACTGATGAGATGTCACACCCAGACGATGGCCA 1990
OY      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyIleLeuIleAlaTyrLeu 80
Db      1991 AACTGGCCGGTAAACCTGGTGGCTGCTGACACCGCGGGGTGGAGCTAATGCGCTATCTC 2050
OY      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db      2051 GGACATTACCAACGACGACGCTGCTGATCCAGGATTCGCGGGAATAAAGGCGTGGCCG 2110
OY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db      2111 CAGTCGTGGTCAGGTCGCGCATTTATGCTGATGTCGACGAGCGCTACCCGTTAAACAGGCGCA 2170
OY      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      2171 GCCTTCGCCCGCGGGCTGGACCTGACTGTCGATGCAATACCGTACATCCAGGAGATACG 2230
OY      141 SerSerProAspProLeuPheAsnProLeuTyrThrGlyValCysGlnLeuAspAsnAla 160
Db      2231 TCACATCCCGATCCGTTATTTAATCTCTTAACAACTGGCGTTGGCACCTGGATACGCG 2290
OY      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      2291 AACGTAGCTGACGCGCATCTCAGCAGGCGAGAGGTCGAATTCGACTTTAACCGGGCAT 2350
OY      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnProGlnSerAsnLeuGly 200
Db      2351 CGGCNAACGGCGTTTCGGAACCTGGAACTGGAAAGGGGCTTAATTTCCGCAATCAACTGTGC 2410
OY      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      2411 CTTAAACGTGAAGAAACAGAGCAAGAAAGCTTTCATTAAACGAGGCAATTACATCGGAATC 2470
OY      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      2471 AAGGTGACCGCGCGCAATGTCTCATTTAACCCGGTGGCTAACCTCGCATCAATGTCGACG 2530
OY      241 GlnIlePheLeuLeuGlnAlaGlnGlnGlyMetProGluProGlyLysArgIleThr 260
Db      2531 GAGATATTCTCTGCAACAAACACACAGGAAATGCCGCGGGGTGGGGAAGCATCC 2590
OY      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db      2591 GATTACACACAGTGAACACCTTGCTAAGTTTGCAATACGCGCAATTTTATTTGCTACAA 2650
OY      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspPheIleMetAlaAla 300
Db      2651 CGCAGCGCAGAGGTGGCCCGCAGCGCCGCAACCCCGTTATTAATGATATCAAGACACCG 2710
OY      301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      2711 TTGAGCCCATCCACCGCAAAACAGCGATGAGTGTCATGATTAACCATCTCACTGCTG 2770
OY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      2771 TTTATCGCGGAGACAGATACATCTGCAATATCGCGCGGCGCACTGAGACTCAACTCGG 2830
OY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      2831 ACGCTTCCCGGTCAAGCCGATTAACACGCGCGCAGGTGTGAACGTGTTGAACGCGCTGG 2890
OY      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      2891 CGTGGCTTAACGATTAACAGCCAGTGATTAAGGTTTGCTGGCTTCCACACTTTACAG 2950
OY      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400

```

```

Db      2951 CAGATGCGTGAATAAACGCCGCTGTCTCATTTAAATACGCCGCCGAGAGGTGAAACTGACC 3010
QY      401  LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
      |||
Db      3011 CTGGCAGAGTGTGAAGAGCCAAATGCGCAGGGCATGTGTCTGTGGCAGGTTTACGCA 3070
QY      421  IleValAsnGluAlaArgIleProAlaCysSerLeu 432
      |||
      3071 ATCGTAATGATGACGACCATACCGCTGCAAGTTTG 3106
RESULT 10
AAC68300
ID      AAC68300 standard; DNA; 17732 BP.
XX
AC      AAC68300;
XX
DT      20-FEB-2001 (first entry)
XX
DE      Lama2/APPA plasmid coding sequence.
XX
KM      Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
XX      environmental pollution; pig; ds.
OS      Mus musculus - chimeric.
XX      Escherichia coli - chimeric.
XX      WO200064247-A1.
XX      02-NOV-2000.
XX      20-APR-2000; 2000WO-CA00430.
XX      23-APR-1999; 99US-0130508.
XX
PA      (UYGU-) UNIV GUELPH.
PI      Forsberg CW, Golovan S, Phillips JP;
XX      WPI: 2000-687245/67.
XX      P-PSDB; AAB36263.
XX
PT      Transgenic non-human animal for gastrointestinal tract specific
PT      expression of a protein, preferably phytase, comprises a nucleic acid
PT      sequence including a heterologous transgene construct encoding the
PT      protein
XX
XX      Claim 14; Fig 23; 152pp; English.
XX
XX      The present invention provides transgenic animals which produce desired
XX      proteins, in this case pigs which expresses phytase in the salivary
XX      gland. Low phytase production levels result in phytate in the diet being
XX      excreted and causing phosphorus contamination in water, as well as
XX      reducing the growth of animals. The invention provides a number of
XX      transgenes containing the E. coli APPA phytase coding sequence.
XX
SQ      Sequence 17732 BP; 4719 A; 4125 C; 4168 G; 4719 T; 1 other:

Alignment Scores:
Pred. NO.:      3.01e-194      Length:      17732
Score:          2235.00      Matches:      430
Percent Similarity: 99.54%      Conservative: 0
Best Local Similarity: 99.54%      Mismatches: 2
Query Match:    97.09%      Indels:      0
DB:            21      Gaps:      0

US-09-777-566a-2 (1-440) x AAC68300 (1-17732)
QY      1  MetLysAlaIleuLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
      |||
Db      12653 ATGAAGCCATCTTAATCCCATTTTATCTCTTGATTCGGTTAAACCCGCAATCTGCA 12712
QY      21  PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
      |||

Db      12713 TTCCGTCAGAGTGAAGCCGAGACCTGAAGCTGGAAGTGTGATTCAGTCATGAT 12772
QY      41  ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
      |||
Db      12773 GTGGCTGCTCCACCAACGAGCCACCACTGATGACGATGTCAACCCACAGACGATGGCCA 12832
QY      61  ThrTrpProValLysLeuGlyTrpLeuThrProArgLysGluLeuIleAlaTrpLeu 80
      |||
Db      12833 ACCTGGCGGTAAGAAACATGGGTGTGGCTGACACCGCGGCTGGTGAATATCGCTACTCTC 12892
QY      81  GlyHisTyrGlnArgGlnArgLeuValAlaAspLysLeuLeuAlaLysGlyCysPro 100
      |||
Db      12893 GGACATTACCAACCCACCGCTGTGTAGCCACGATTCGGGAAAGGGCTGCCCG 12952
QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgTrpArgLysThrGlyGlu 120
      |||
Db      12953 CAGTGTGTCAGGTGCGATTATGTGATGTGACGAGCGGTACCGCTAAACAGGGGAA 13012
QY      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
      |||
Db      13013 GCCTTCGCGCGCGGCTGGCACCCTGACTGTGCATTAACCTACTACCCAGCGCATACG 13072
QY      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
      |||
Db      13073 TCCAGTCCCGATCCGTTATTAAATCTTAATAAACTGGCTTCCCAACTGGATAACGCG 13132
QY      161 AsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGlyHis 180
      |||
Db      13133 AACGTGACGACGCGATCTCAGCAGGAGGAGGCTCAATGCTGACTTACCGGGCAT 13192
QY      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
      |||
Db      13193 CGGCAACGGCGTTTCGCGAATCGAAGCGGTGCTTATTTCCGCATCAATCAATTGTGCG 13252
QY      201 LeuLysArgGluLysGlnAspLysCysSerLeuThrGlnAlaLeuProSerGluLeu 220
      |||
Db      13253 CTTAAAGGTGAGAAACAGAGAAAGCTGTTCATTAACGCGCATTAACATCGGAACCTC 13312
QY      221 LysValSerAlaAspAsnValSerLeuThrGlyValAlaSerLeuAlaSerMetLeuThr 240
      |||
Db      13313 AAGGTGACCGCGCAATGTCTATTACCGGTGCGGTACCTCGCATCATCTGACG 13372
QY      241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
      |||
Db      13373 GAGATATTTCCTCCGCAACAAGACAGAGGAATGCCGAGCGGGGTGGGAGATCAACC 13432
QY      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
      |||
Db      13433 GATTTCACACACAGTGGAAACACCTTCCTAAAGTTTGCATTAACGCGCAATTTTATTTGCTACAA 13492
QY      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
      |||
Db      13493 CGCAGCGCAGAGGTTGCCCGGAGCGCGCCACCCGCTTATTAAGATTGATCAACAGACGCG 13552
QY      301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
      |||
Db      13553 TTGAGCGCCCATTCACACGCAAAACAGCGCATGTGTGACATTAACCACTCACTGCTGCTG 13612
QY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
      |||
Db      13613 TTTATCCCGGACACGATTAATCTGCAAAATCTCGCGCGCCGACGAGCTCAACTGAG 13672
QY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyLysGluLeuValPheGluArgTrp 360
      |||
Db      13673 ACGCTTCCCGGTCAAGCGGATTAACACGCGCAGGTGTAAGTGTGTTGAACGCTGG 13732
QY      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
      |||
Db      13733 CGTGGCTAAGCGATTAACAGCAGCACTGATTCAGTTTCGCGTCTCCAGACTTTACAG 13792
QY      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
      |||
Db      13793 CAGATGCGTGAATAAACGCCGCTGTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 13852

```

QY 401 LeuAlaGlyCysGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGAGAGATGTGAAGACGAAATGCGCAGGCGATGTGTGGTTCGAGGTTTACGCA 13912
 QY 421 IleValaGlnAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGAATGAAGCAGCATACCGCTTGACAGTTTG 13948
 RESULT 11
 AAC68294
 ID AAC68294 standard: DNA; 20623 BP.
 XX AAC68294;
 AC AAC68294;
 XX 20-FEB-2001 (first entry)
 DT 20-FEB-2001 (first entry)
 XX
 DE Lama2/APPA plasmid coding sequence.
 XX
 KM Transgenic animal; salivary protein; phytase; phosphorus; animal growth;
 KM environmental pollution; pig; ds.
 XX
 OS Mus musculus - chimeric.
 OS Escherichia coli - chimeric.
 XX
 PN WO200064247-A1.
 PD 02-NOV-2000.
 PE 20-APR-2000; 2000WO-CA00430.
 PR 23-APR-1999; 99US-0130508.
 PA (UYGU) UNIV GUELPH.
 PI Forsberg CM, Golovan S, Phillips JP;
 XX
 DR WPI: 2000-687245/67.
 DR P-PSDB: AAB36257.
 PT Transgenic non-human animal for gastrointestinal tract specific
 PT expression of a protein, preferably phytase, comprises a nucleic acid
 PT sequence including a heterologous transgene construct encoding the
 PT protein -
 XX
 PS Claim 56; Fig 5; 152pp; English.
 CC The present invention provides transgenic animals which produce desired
 CC proteins, in this case pigs which expresses phytase in the salivary
 CC gland. Low phytase production levels result in phytate in the diet being
 CC excreted and causing phosphorus contamination in water, as well as
 CC reducing the growth of animals. The invention provides a number of
 CC transgenes containing the E. coli APPA phytase coding sequence.
 XX
 SQ Sequence 20623 BP; 5449 A; 4847 C; 4902 G; 5424 T; 1 other;
 Alignment Scores:
 Pred. No.: 3,71e-194 Length: 20623
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 21 Gaps: 0
 US-09-777-566a-2 (1-440) x AAC68294 (1-20623)
 QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 12653 ATGAAGGCATCTAATCCATTTTATCTTCGATTCGTTAAACCCCAATCTGCA 12712
 QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
 DB 12713 TTCGCTCAGATGAGCCGAGACTGAGAGTGGAAAGTGGATTTGTCAGTGTCATGCT 12772

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 DB 12773 GTGGTGTCTCCAAACCAAGCCACGCACTGATGAGATGTCACCCAGACGCAATGGCA 12832
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGlnLeuIleAlaTrpLeu 80
 DB 12833 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGGTGGTGAATGATGCGCTATCTC 12892
 QY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
 DB 12893 GGACATTCACCAACCGCAGCTGCTGTAACCCAGGATTCGTGGCAAAAAGCGCTGCCG 12952
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 12953 CAGTCTGTCAGGTGCGCATATTATTCGATGTCGACGAGCGTACCCGTAACAGCGCA 13012
 QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 13013 GCCTTCGCGCGCGGCTGGCACCTGACTGTCATTAACCGTAATACCCAGCAGATPAG 13072
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 13073 TTCAGTCCCGATCCGTTTATTAATCTTAAAACTGGCGTTTGCCAACTGATPACGCG 13132
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
 DB 13133 AACGTACTGACGCGATCTCAGCAGGCGAGAGGTCATATGCTGACTTACGGCGCAT 13192
 QY 181 ArgGlnThrAlaPheArgGluLeuGlnLysValLeuAsnProProGlnSerAsnLysCys 200
 DB 13193 CGGCAAACGGCGTTCGCGAACTGGAAAGGGGTCTTAATTTCCGCAATCAAACTGTGCG 13252
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 13253 CTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGCAATTCATTCGGAATCC 13312
 QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 13313 AAGGTAGCGCGCGCAATGCTCATTAACCGGTGCGTAAACCTCGCATCAATGCTGACG 13372
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGlyTrpGlyArgIleThr 260
 DB 13373 GAGATATTTCTCTGCAACAGCACAGGGAATGCGGAGCGGGGTGGGAAGGTACGCC 13432
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
 DB 13433 GATTACACCACTGGAACACCTGCTAGTTTGCATACGGCGCAATTTTATTTGCTACAA 13492
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 13493 CGCAGCGCAGAGGTGCGCGCAGCGCGCCACCCCGTATTAAGATTATCATCAGACACGCG 13552
 QY 301 LeuThrProHisProProGlnLysGlnAlaArgValIleThrLeuProThrSerValLeu 320
 DB 13553 TTGACGCCCATCCACCCCAAAACAGCGTATGTTGATGATTAACCTTACCTTACGTGCTG 13612
 QY 321 PheIleAlaGlyHisAspThrAsnLeuValAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
 DB 13613 TTTATCGCGGAGACAGATACATTAATCTGCAATCTCGCGGCGGCACTGAGACTCAACTGG 13672
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
 DB 13673 ACGCTTCCCGGCTCAGCCGATTAACACGCGCGCAGTGTGTAACGTGTTTGAACGCGTGG 13732
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
 DB 13733 CGTGGCTTAACCGAATACAGCAGAGGATTAAGTTTGGCTGCTTCCAACTTACAG 13792
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyGluValLysLeuThr 400
 DB 13793 CAGATGCGTGAATAAACGCCGCTGTCAATTAATACGCCGCGGAGAGAGTAAACTGACC 13852

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 DB 13853 CTGGCAGAGATGTAAAGAGCAATGCGCAGGCGATGTGTCTGGCAGGTTTACGCA 13912

QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
 DB 13913 ATCGTGAATGACACGACATACCCGCTTGCGAGTTTG 13948

RESULT 12
 AAA28216
 ID AAA28216 standard; DNA: 1489 BP.
 AC AAA28216;
 XX
 DT 12-FEB-2001 (first entry)
 XX
 DE E. coli acidic phosphatase appa2 encoding nucleotide sequence.
 XX
 KW Phosphatase; phytase activity; protease; acidic phosphatase; appa2;
 KM periplasmic phosphonhydrolase phosphonhydrolase; animal feed; ds.
 XX
 OS Escherichia coli.
 XX
 FH Key location/Qualifiers
 FH protein_bind complement (1..22)
 FT /*tag- a
 FT /bound_moiety- "Primer Pfl"
 FT 16..108
 FT /*tag- b
 FT /product- "Appa2 related peptide"
 FT 184..1482
 FT /*tag- c
 FT /product- "Appa2"
 FT /*note- "acidic phosphatase"
 FT 241..261
 FT /*tag- d
 FT /bound_moiety- "Primer E2"
 FT 1471..1489
 FT /*tag- e
 FT /bound_moiety- "Primer K2"
 XX
 PN W0200058481-A2.
 PD 05-OCT-2000.
 PF 31-MAR-2000; 2000MO-US08590.
 XX
 PR 31-MAR-1999; 99US-0127032.
 XX
 PA (CORR) CORNELL RES FOUNDD INC.
 PI Let X;
 DR WPI: 2000-619081/59.
 DR P-PSDB: AAY94753, AAY94756.
 XX
 PT Novel phosphatase having improved phytase activity useful as animal
 PT feed for improving the accessibility of phosphate to an animal is
 PT produced by treating phosphatase with protease
 XX
 PS Disclosure; Fig 6; 48pp; English.
 XX
 CC This invention relates to a phosphatase fragment with improved phytase
 CC activity. The phosphatase fragment is created by treating the phosphatase
 CC with a protease. The invention includes an Escherichia coli nucleotide
 CC sequence which encodes an acidic phosphatase appa2 which has improved
 CC phosphatase activity. The appa gene of Escherichia coli was originally
 CC defined as E. coli periplasmic phosphonhydrolase phosphonhydrolase,
 CC although it was not previously known to have phytase activity. The
 CC phosphatase fragment can be used in animal feed, to improve the
 CC accessibility of phosphate to the animal. The phosphatase gene which
 CC encodes the protein can be used in the production of phytase. The present
 CC sequence represents the appa2 coding sequence.

XX
 SQ Sequence 1489 BP; 362 A; 384 C; 403 G; 340 T; 0 other;
 Alignment Scores:
 Pred. No.: 1,9e-194 Length: 1489
 Score: 2221.00 Matches: 428
 Percent Similarity: 99.07% Conservative: 0
 Best Local Similarity: 99.07% Mismatches: 4
 Query Match: 96.48% Indels: 0
 DB: 21 Gaps: 0

US-09-777-566A-2 (1-440) x AAA28216 (1-1489)

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
 DB 182 ATGAAAGCGATCTTAATCCATTTTATCTTTGATTCCTTAACCCCGCAATCTGCA 241

QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 DB 242 TTGCTCAGAGTAGCGCGAGCTGAAGCTGGAAGTGTGTGATTTGTCACCGCTCATGCT 301

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
 DB 302 GTGCGTGCCCCCAACCAAGCCACGCAACTGATGACAGATGTACCCCGACGATGGCCA 361

QY 61 ThrTriProValLysLeuGlyTTPleuThrProArgLysGlyGluLeuIleAlaTyrLeu 80
 DB 362 ACCTGCGCGGTAAACGTGGTGTGCTGACACCGCGGTGTAGCTTAATCGCTATCTC 421

QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
 DB 422 GGACATTACCAACGCGCGCTGTGTGGCCGACGAGATGTGGCGAAAGGGCTGCGCG 481

QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
 DB 482 CAGCTGTGTCAGGTGCGGATTAATGATGATGACAGCGCTACCCGTAAACAGCGGAA 541

QY 121 AlaPheAlaIleGlyLeuAlaIleProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 DB 542 GCCCTGCGCGCGCGCGCGCTGACTGTGCATTAACCTGATACCGGAGAGATACG 601

QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 DB 602 TCCAGTCCGATCCGTTATTTATTCCTTAATAAAGTGGCTTGCCAACTGATTAACGCG 661

QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 DB 662 AACGTGACTAGCGGATCTCAGCAGGCGAGGAGGTCAATTGCTGACTTACCGGGCAT 721

QY 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
 DB 722 CGGCAAAAGCGGCTTTCGCGAAGTGAAGCGGCTCTTAATTTTCCCAATTAACCTGTGC 781

QY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 DB 782 CTTAACCGCTAGAAACAGCAAGCAAGAGTGTTCATTAAACGACGATTAACATGCAACTC 841

QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 DB 842 AAGGTGAGCCCGCAACATGTTTCATTAAACGCGGTGTAAGCTCCGATCAATGAAGTGCAGC 901

QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTrrpGlyArgIleThr 260
 DB 902 GAATATATTCTCTGCAACAGCAAGCAAGAGTAATCCGAGCGGGGTGGGAAGATCACT 961

QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
 DB 962 GATTACACCAAGGGAACACTTGTCTAAGTTTGATTAACCGCAATTTTATTACTACAA 1021

QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 DB 1022 CGCAGCGCAAGAGGTGGCCGCAAGTGGCGCACCCCGCTTATTTGATTGATTCAGGAGCG 1081

Dd	668	AACGTGATGACCGGATCTCAGCAGGCGACAGGGTCAATTGCTGACTTTACCGGGCAT	727
Qy	181	ArgcInthralaAtheatrgcInleuGluArgValleuasnphroGlnSerAsnleuGys	200
Dd	728	CGGCAAAACGGCGTTCGCGAACTGGAAACGGGCTGCTTAATTTTCCGCAATCAAACTTGTC	787
Qy	201	LeuysArgrgluGysGlnAspGluSerCysSerleuthrglnAlaleuProSerGluLeu	220
Dd	788	CTTAAACCTGGAAACACGAGACGAAAGCTGTTCAATTAAACGAGGCACTTACCAATCGGAATCTC	847
Qy	221	LysValSerAlaAspAsnValSerleuthrglAlaValSerleuAlaSerMetleuthr	240
Dd	848	AAGGTGACCGCCGCAATGTTTCATTTACCGGATGGCGTAAGCTCGCATCAATGCTGACG	907
Qy	241	GluIlePheIleuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThr	260
Dd	908	GAATATATTTCTCCGCAACAAAGCACAGGGAATGCCGAGGCCGGGTGGGGAAGATCCACT	967
Qy	261	AspSerHisGlnTyrPasnThrleuLeuSerleuHisAsnAlaGlnPheTyrleuLeuGln	280
Dd	968	GATTCACACACAGTGGAAACACCTTGTTGTAAGTTTGATACCGGCATTTATTTACTACAA	1027
Qy	281	ArgThrProGluValAlaArgSerIrgAlaThrProleuLeuAspIleuMetAlaAla	300
Dd	1028	CGCACGCCAGAGGTTGCCCGGACAGTCGCGCACCCCGCTTATTTGATTCATACAGACACCG	1087
Qy	301	LeuthrProHisProProGlnLysGlnAlaTyrGlyValIthrleuProThrSerValleu	320
Dd	1088	TTTGAGCGCCCATCCACCGCAAAACAGCGCTATGTGTGACATTTACCACTTCAGTCTG	1147
Qy	321	PheIleAlaGlyHisAspThrAsnleuAlaAsnleuGlyGlyAlaLeuGluLeuAsnTrp	340
Dd	1148	TTTATTTGCGGACACGATACTTAATCTGCGCAATCTCGGCGCGCACTGAGCTCAACTGG	1207
Qy	341	ThreupProGlyGlnProAspAsnThrProProGlyGlyGluIleuValPheGluArgTrp	360
Dd	1208	ACGCTTCAGGTCCACCGGATTAACCGCGCGCAAGGTGTGAACCTGTGTTGAACGCTGG	1267
Qy	361	ArgThrleuSerAspAsnSerGlnTrpIleGlnValSerleuValPheGlnIthrleuGln	380
Dd	1268	CGTGGGCTAAACGATTAACAGCCAGTGAATTCAGGTTTGGTGCTTCACAGCTTACAG	1327
Qy	381	GlnMetArgAspIlysthrProleuSerleuAsnThrProProGlyGluValLysleuthr	400
Dd	1328	CAGATGGCGTGAATAAACCCGCTATCATTTAATACGCCGCCGAGAGTGAAACTGAC	1387
Qy	401	LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerleuAlaGlyPheThrGln	420
Dd	1388	CTGGCAGAGATGTGAAGACGGAATGCGCAGGGCATGTGTTCGTGGCCGGTTTACGCCAA	1447
Qy	421	IleValAsnGluAlaArgIleProAlaCysSerleu 432	
Dd	1448	ATCGTGAATGAAGCGCGCATACCGCGCTGCAGTGTG 1483	
RESULT 14			
ID	AAC68299	standard; DNA; 5421 BP.	
XX	AAC68299;		
XX	20-FEB-2001	(first entry)	
DE	SV40/APPA plasmid coding sequence.		
XX	Transgenic animal; salivary protein; phylase; phosphorus; animal growth;		
KW	environmental pollution; pig; ds.		
OS	Rhesus macaque polyoma virus - chimeric.		
OS	Escherichia coli - chimeric.		
XX	MO200064247-Al.		
XX	02-NOV-2000.		
DD			

XX	20-APR-2000; 2000OWO-CA00430.
PF	
XX	23-APR-1999; 99US-0130508.
PR	
XX	(UYCU-) UNIV GUELPH.
PA	
XX	Forsberg CW, Golovan S, Phillips JP;
P1	
XX	WPI; 2000-687245/67.
DR	
XX	P-PSDB; AAB36262.
XX	
PT	Transgenic non-human animal for gastrointestinal tract specific
PT	expression of a protein, preferably phytase, comprising a nucleic acid
PT	sequence including a heterologous transgene construct encoding the
XX	protein
PS	
XX	Claim 56; Fig 22; 152pp: English.
CC	
XX	The present invention provides transgenic animals which produce desired
CC	proteins, in this case pigs which expresses phytase in the salivary
CC	gland. Low phytase production levels result in phytate in the diet being
CC	excreted and causing phosphorus contamination in water, as well as
CC	reducing the growth of animals. The invention provides a number of
CC	transgenes containing the E. coli APPA phytase coding sequence.
XX	
SQ	Sequence 5421 BP; 1413 A; 1321 C; 1331 G; 1355 T; 1 other;
..	
Alignment Scores:	
Pred. No.:	3,35e-192 Length: 5421
Score:	2,205.00 Matches: 429
Percent Similarity:	99.31% Conservative: 0
Best Local Similarity:	99.31% Mismatches: 3
Query Match:	95.79% Indels: 1
DB:	Gaps: 0
US-09-777-566A-2 (1-440) x AAC68299 (1-5421)	
OY	1 MetlysalalileulleleProPheLeuSerLeuLeulleProLeuthrProGlnSerAla 20
Db	40 ATGAAGGCATCTTAACCCATTTTTATCTCTTGATTCCGTAAACC CGCATTCGCA 99
OY	21 PhealaGlnSerGIUProGlnLeuLysLeuISerValVallIleValISerArgHisGly 40
Db	100 TTCGCTCAGATGTAGCGCGAGCTGAAGCTGGAAAGTGTGGTATTGTACATCCATCATGT 159
OY	41 ValArGlaAProThrnrllysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db	160 GTGGGTGTCTCAACCAAGCCAGCCACCACTGATGTGAGATGTACACC CCAGACGATGGCCA 219
OY	61 ThrTrpProValLysLeuGLYTrpLeuThrProArgGlyGlyGluLeuIleAlaTyLeu 80
Db	220 ACCTGCGCCGATAAACCTGGTGGCTGTGACCCGCCNGGTGGTAGCTAAATCGCTATCTC 279
OY	81 GlyHstYrGlnArGInArGInArGLeuValAlaAspGlyLeuLeuAlaLysLysGlyCysPro 100
Db	280 GGACATTPACCAAGCCAGCGCTGTGTAGCCAGACGATGTGGCGAA AAAAGGCGTCCCG 339
OY	101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db	340 CAGTCTGTGTCAAGGTGCGCATTTATGTCGATGTCAGACGAGCGGTACCCCTAAAACAGCGGAA 399
OY	121 AlaPheaIleAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db	400 GCCTTCGCGCGCGCGGCTGGCACCTGTACTGTGCATTAACCGTACTATACCCAGGCAAGATACG 459
OY	141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db	460 TCACGTCCCGCATCGGTATTTAAATTCCTCTAAAAACTGGCGTTTCCCAATCGAATACGCG 519
OY	161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db	520 AACGGATCTGACGGCATCTTCAGCAGGAGGAGGATCAATTGCTGACTTTACCGGCGCAT 579

|||||
Db GTGGGTGCCCCAACCAAGGCCACCACTGATGCGAGATGTCACCCGAGCATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrLeu 80
Db ACCGTGGCCGGTAAACATGGGTGGTGGTGCACACACCGGTGGTGGTGAATGCTATGCTC 427
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuAlaLysGlyCysPro 100
Db 428 GGAATATACCAACCCACCGCTGTGGTGGCCACGATGGCTGGGAAAAAGGCTGCCCG 487
QY 101 GlnSerGlyValAlaIleIleAlaAspValAspGlnArgTyrGlyGlu 120
Db 488 CAGCGTGTGTCAGTCGCGATATGTCGATGTCAGCAGCGTACCCGTAACAAACAGCGCA 547
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCGTTCGCGCGGGGCTGGCCACTGCTGCAATACCGTACATACCCAGCAGATACG 607
QY 141 SerSerProAspProLeuPheAsnProLeuTyrThrGlyValCysGlnLeuAspAsnAla 160
Db 608 TCCAGTCCCGATCCGTTATTTATCTCTTAACAACTGGCGTTGGCACTGATTAACGCG 667
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTGACTGACGGGATCTCAGCAGGAGGAGGATCAATGCTGACTTACCGGCGCAT 727
QY 181 ArgGlnThrAlaPheArgGlyLeuGlnArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCAAAACGGGTTCCGAACTGAAACGGGTTGCTTAATTTCCGCAATCAAACTTGAC 787
QY 201 LeuTyrArgGlyLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 788 CTTAAACGTGAAACAAAGATGAAGCTGTAACTTAACGACGATTAACGAACTC 847
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTGAGCGCGCAATGTTTCATTAACGCGTGGTAAGCTCGCATCATGCTGACG 907
QY 241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTyrGlyArgIleThr 260
Db 908 GAAATATTTCTCTGCAACAAACAGCAGGAAATGCCGGGCGGGTGGGGAAGATCACT 967
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 968 GATTACACACAGTGGAAACACTTCTTAAGTTTGCATTAACGCGCAATTTTATTCTACAA 1027
QY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 1028 CGCAGCGCAGAGTGGCCGCGAGTCGCGCCACCCGTTATGATTCAGATTCAGACGCG 1087
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGACGGCCCATCCACCGCAAAACAGCGCATGTGTGACATTAACCCACTTCAGTCTG 1147
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTrp 340
Db 1148 TTTTATTCGCGACACGATCTAATCTGCGCAATCTCGCGGCGCATGAGCTCAACTGCG 1207
QY 341 ThrLeuProGlnProAspAsnThrProProGlyGlyLeuValPheGlnArgTrp 360
Db 1208 ACGCTTCAGGTACGCGGATTAACACGCGCCAGGTGGTGAACGTGTTGAACGCTGCG 1267
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1268 CGTGGGTAAACGATACAGCAGTGAATTCAGGTTTCGCTGCTCCAGACTTTACAG 1327
QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnValLysLeuThr 400
Db 1328 CAGATGCGGTAAACGCGCTCTCATTAATACGCGCGCGGAGAGGTGAACCTGACC 1387
QY 401 LeuAlaGlyCysGlnArgGlnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
|||||

Db 1388 CTGGCAGAGTGTGAAGACGAATGCGCAGGCGCATGTGCTGTGGCCGTTTACGCA 1447
QY 421 IleValAsnGlnAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTAATGAAGCGCGCATACCGCGCTGCAAGTTTG 1483

Search completed: June 12, 2003, 13:27:26
Job time : 408 secs

THIS PAGE BLANK (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 12:11:14 ; Search time 1929 Seconds

(without alignments)
3694.149 Million cell updates/sec

Title: US-09-777-566A-2

Sequence: 1 MKAILPLSLILPLTPQSA.....IVNEARIPACSLRHHHHH 440

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO/spool/US09777366/runat_05062003_154808_6562/app_query.fasta_1.583
-DB=EST -QFMT=fastap -SUFFIX=1st -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi -LIST=45
-OCALIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=fto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09777366 -CGC=1.1 1996 -etunat_05062003_154808_6582 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEDUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV.TIMEOUT=120 -WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_estlum:*
3: em_estlin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estlum:*
16: em_estlum:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rnd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	104.4	45.4	616	13	Bj074127	Bj074127 Bj074127
C 2	64.3	27.9	383	10	AM036132	AM036132 EST274508
C 3	64.3	27.9	383	10	AM036134	AM036134 EST274510
C 4	64.1	27.8	393	10	BES20240	BES20240 M1B125TM
C 5	511.5	22.2	354	9	AA545747	AA545747 HBMSP1B4 -
C 6	392.7	17.0	320	10	BEA36403	BEA36403 EST407481
C 7	355	15.4	208	13	BM409016	BM409016 EST183343
C 8	355	15.4	208	13	BM412806	BM412806 EST587133
C 9	339	14.7	546	13	Bj040796	Bj040796 BJ040796
C 10	287	12.5	657	12	BG457132	BG457132 NF100C05P
C 11	278	12.1	179	10	BE461872	BE461872 EST413291
C 12	269	11.7	508	13	Bj030203	Bj030203 Bj030203
C 13	228.5	9.9	480	9	AL586973	AL586973 AL586973
C 14	225	9.8	254	10	BES20241	BES20241 M1B12XTM
C 15	172.5	7.5	696	9	AU004475	AU004475 AU004475
C 16	144.5	6.3	645	13	Bj037937	Bj037937 Bj037937
C 17	125	5.4	274	14	B0153360	B0153360 NF035C07I
C 18	119.5	5.2	557	13	Bj029368	Bj029368 Bj029368
C 19	118	5.1	1075	14	BM922358	BM922358 AGENCOURT
C 20	116.5	5.1	720	13	BM628471	BM628471 170006874
C 21	114.5	5.0	702	12	BF037418	BF037418 601460789
C 22	114	5.0	1010	13	BM465775	BM465775 AGENCOURT
C 23	112.5	4.9	1060	14	BM922665	BM922665 AGENCOURT
C 24	112	4.9	1797	11	BC013495	BC013495 Mus muscu
C 25	109.5	4.8	1080	11	BQ278048	BQ278048 AGENCOURT
C 26	109	4.7	2396	11	BC022504	BC022504 Homo sapi
C 27	108.5	4.7	766	13	B1764510	B1764510 603050802
C 28	108	4.7	695	10	AV400670	AV400670 AV400670
C 29	107.5	4.7	657	13	B1265620	B1265620 NF096B07I
C 30	107	4.6	686	12	BG457966	BG457966 NF036A10P
C 31	107	4.6	842	12	BG026657	BG026657 602293790
C 32	107	4.6	1114	14	BM920972	BM920972 AGENCOURT
C 33	105.5	4.6	915	9	AL545650	AL545650 AL545650
C 34	105	4.6	567	10	AM666304	AM666304 SK34H10.Y
C 35	104.5	4.5	1129	14	BM926193	BM926193 AGENCOURT
C 36	104.5	4.5	1212	13	BM547032	BM547032 AGENCOURT
C 37	104	4.5	565	13	BM084995	BM084995 sa350h07
C 38	103.5	4.5	775	13	B1819484	B1819484 603034283
C 39	103	4.5	548	12	BF519356	BF519356 EST456818
C 40	103	4.5	924	12	BE889074	BE889074 601513513
C 41	102.5	4.5	1907	12	BF166252	BF166252 601770114
C 42	102	4.4	621	10	AM175388	AM175388 f135e03.Y
C 43	102	4.4	1042	14	BM903520	BM903520 AGENCOURT
C 44	101.5	4.4	1161	14	BM926491	BM926491 AGENCOURT
C 45	101.5	4.4	1197	13	B1489085	B1489085 603021120

ALIGNMENTS

RESULT 1
Bj074127/c
LOCUS Bj074127 616 bp mRNA linear EST 11-DEC-2001
DEFINITION laevis CDNA clone X1090106 5', mRNA sequence.
ACCESSION Bj074127
VERSION Bj074127.1 GI:17504316
KEYWORDS EST.
SOURCE African clawed frog.
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 616)
Kitayama,A., Terasaka,C., Mochii,M., Ueno,N., Shin-I,T. and Kohara

TITLE
JOURNAL
COMMENT

Expressed genes in X. laevis embryo
Unpublished (2001)
Contact: Tadashi Shin-I
Center For Genetic Resource Information
National Institute of Genetics
1111 Yata, Mishima, Shizuoka 411-8540, Japan
Tel: 81-559-81-6856
Fax: 81-559-81-6855
Email: tshin@genes.nig.ac.jp.

FEATURES
source

Location/Qualifiers
1.616
/organism="Xenopus laevis"
/db_xref="taxon:8335"
/clone="X1090106"
/clone_lib="NIBB Mochi normalized Xenopus tailbud library"
/tissue_type="whole embryo"
/dev_stage="stage 25"
/note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

BASE COUNT 128 a 167 c 167 g 154 t
ORIGIN

Alignment Scores:

Pred. No.: 7.59e-100 length: 616
Score: 1044.00 Matches: 204
Percent Similarity: 99.51% Conservative: 0
Best Local Similarity: 99.51% Mismatches: 1
Query Match: 45.35% Indels: 0
DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BJ074127 (1-616)

QY 89 ValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIle 108
DB 615 GTAGCCGAGGATGCTGCGGAAAGGCGCTCCGACGCTGTCAGTCGCGGATATT 556
QY 109 AlaAspValAspGlnArgThrArgLysThrGlyGlnAlaPheAlaIleGlyLeuAlaPro 128
DB 555 GCTGATGTCGACGAGCGTACCGCTAAACGCGGAAAGCTTCCGCCCGGGCTGGACCT 496
QY 129 AspCysAlaIleThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsn 148
DB 495 GACTGTGCATATACGTACATACCGACGAGATACGTCGATCCGATCCGATTATTAA 436
QY 149 ProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSer 168
DB 435 CCTCTAAAACCTGGCTTGGCCACACGATTAACGCAACGTCGATGACGCGATCTCAGC 376
QY 169 ArgAlaGlyGlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGlnLeu 188
DB 375 AGGCGAGAGGAGCAATGCTGACTTACCGGCGCATGGCAAAAGCGCTTCCGGAATCG 316
QY 189 GluArgValLeuAsnProGlnSerAsnLeuCysLeuLysArgGlnLysGlnAspGln 208
DB 315 GAACGGGCTCTTAATTTCCGCAATCAACTTGTGCTTAAAGTGAGAAACAGACGAA 256
QY 209 SerCysSerLeuThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspAsnValSer 228
DB 255 AGCTGTTCATTAAACGAGCATATACATCATGGAAGTGAAGCCCGCAATGTCTCA 196
QY 229 LeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuGlnGlnAla 248
DB 195 TTAAACGGGTGGCTAAAGCTCGCATCAATGCTGACGAGATATTCTTCCGCAACAGCA 136
QY 249 GlnGlyMetProGlnProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeu 268
DB 135 CTGGGAATGCGGACCGGGGTGGGAGAGATACCGGATTCACACCACTGGAAACACTTG 76

QY 269 LeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGlnArgThrProGlnValAlaArgSer 288
DB 75 CTAAAGTTGATTAACGCGCAATTTATTCTCAACAGCAGCGGAGTTGCCGACG 16
QY 289 ArgAlaThrProLeu 293
DB 15 CGCGCACCCCGTTA 1

RESULT 2
AM036132 383 bp mRNA linear EST 18-MAY-2001
LOCUS
DEFINITION
EST274508 tomato seed, TAMU Lycopersicon esculentum cDNA clone
cLEIE23 similar to periplasmic phosphonhydride phosphohydrolase
precursor, putative, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

AM036132
EST
tomato.
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE
AUTHORS
1 (bases 1 to 383)
Alcala,J., Vredalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
Liang,F., Upton,J., Romning,C.M., Craven,M.B., Fujii,C.Y., Bowman
C.L., Nierman,M., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
S.D. and Giovannoni,J.
Generation of ESTs from tomato seed tissue
Unpublished (1999)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
3 prime sequence.

FEATURES
source

Location/Qualifiers
1.383
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEIE23"
/clone_lib="tomato seed, TAMU"
/tissue_type="seeds"
/dev_stage="quiescent seed"
/lab_host="X11-Blue MRF"
/note="Vector: Bluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEIE - Tomato seed EST library. Directionally cloned
cDNAs inserted into Bluescript SK(-) at 5' end with
EcoRI and 3' end with XhoI site."

BASE COUNT 102 a 101 c 97 g 83 t
ORIGIN
Alignment Scores:
Pred. No.: 8.28e-58 length: 383
Score: 643.00 Matches: 127
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 27.93% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x AM036132 (1-383)

QY 117 LysThrGlyGlnAlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThr 136
DB 2 AAAACAGCGCAACCTTCCGCCCGGCTGGACCTGACGTGTCATTAACCGTACATACC 61
QY 137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGln 156
DB 62 CAGGAGATACGTCACAGTCCGATTCGATTAATTCCTTAATAACTGGCGTTGGCAA 121
QY 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaIleGlySerIleAlaAsp 176

Db 122 CTGATTAACGCCAGCTGACGCGATCTCAGACGAGGAGAGGTCATATGCTGCAC 181
 Oy PheThrglYhISArGlnThrAlaPheArgIuLeuGluArgValLeuAsnProGln 196
 Db 177 TTTACCGGCGATCGCAACGCGGCTTCCGCAACGTGAAACGCGGCTTAAATTTCCGCAA 241
 Oy 182 TTTACCGGCGATCGCAACGCGGCTTCCGCAACGTGAAACGCGGCTTAAATTTCCGCAA 241
 Db 197 SerAsnLeuCySLeulysArgIuLysGlnAspGluSerCySerLeuThrglAlaLeu 216
 242 TCAAACTGTGCTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGATTA 301
 Oy 217 ProSerGluLeuLysValSerAlaAspAsnValSerLeuThrglAlaValSerLeuAla 236
 Db 302 CCATCGGAACCTCAAGCTGAGCGCCGACATGTCATTAACGCGGTGTAAGCTCGCA 361
 Oy 237 SerMetLeuThrglIlePhe 243
 Db 362 TCAATGCTGACGAGATATTT 382
 RESULT 3
 AM036134/c
 LOCUS 383 bp mRNA linear EST 18-MAY-2001
 DEFINITION ES2724510 tomato seed, TAMU Lycopersicon esculentum cDNA clone
 cLE1E23 similar to periplasmic phosphonhydrolase precursor, putative, mRNA sequence.
 ACCESSION AM036134
 VERSION AM036134.1 GI:5894813
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
 Lycopersicon.
 1 (bases 1 to 383)
 Alcalá, J., Vrebalov, J., White, R., Matern, A. L., Vision, T., Holt, I. E.,
 Liang, F., Upton, J., Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman,
 C. L., Nierman, W., Fraser, C. M., Venter, J. C., Martin, G. B., Tanksley,
 S. D. and Giovannoni, J.
 Generation of ESTs from tomato seed tissue
 Unpublished (1999)
 TITLE JOURNAL
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.
 FEATURES
 source
 1..383
 /organism="Lycopersicon esculentum"
 /cultivar="T496"
 /db_xref="taxon:4081"
 /clone="cLE1E23"
 /clone_lib="tomato seed, TAMU"
 /tissue_type="seeds"
 /dev_stage="quiescent seed"
 /lab_host="XLI-Blue MRF"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI, Site 2:
 XhoI; cLE2 - Tomato seed EST library. Directionally cloned
 cDNAs inserted into pBluescript SK(-) at 5' end with
 EcoRI and 3' end with XhoI site."
 BASE COUNT 83 a 97 c 101 g 102 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 8,28e-58 Length: 383
 Score: 643.00 Matches: 127
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 27.93% Indels: 0
 DB: 10 Gaps: 0
 US-09-777-566a-2 (1-440) x AM036134 (1-383)

Oy 117 LysThrglYgluAlaPheAlaIleuAlaProAspCyAlaIleThrValHisThr 136
 Db 382 AAAACAGCGGAAGCCTTCCCGCGCTGCGACCTGACTGTGCATTAACCTGATACC 323
 Oy 137 GlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrglYalCysGln 156
 Db 322 CAGCGAGATACGCGACGCTCCGATCCGTTATTTAATCCTTAATAAACTGGCTTGGCCA 263
 Oy 157 LeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAsp 176
 Db 262 CTGATTAACGCAACCTGACTGACGATCTCAGCAGGAGAGAGGTCATATGCTGCAC 203
 Oy 177 PheThrglYhISArGlnThrAlaPheArgIuLeuGluArgValLeuAsnProGln 196
 Db 202 TTTACCGGCGATCGCAACGCGGCTTCCGCAACGTGAAACGCGGCTTAAATTTCCGCAA 143
 Oy 197 SerAsnLeuCySLeulysArgIuLysGlnAspGluSerCySerLeuThrglAlaLeu 216
 Db 142 TCAAACTGTGCTTAAACGTGAGAAACAGACGAAAGCTGTTCAATTAACGACGATTA 83
 Oy 217 ProSerGluLeuLysValSerAlaAspAsnValSerLeuThrglAlaValSerLeuAla 236
 Db 82 CCATCGGAACCTCAAGCTGAGCGCCGACATGTCATTAACGCGGTGTAAGCTCGCA 23
 Oy 237 SerMetLeuThrglIlePhe 243
 Db 22 TCAATGCTGACGAGATATTT 2
 RESULT 4
 BE520240/c
 LOCUS 393 bp mRNA linear EST 19-MAR-2001
 DEFINITION M1B125M Arabidopsis developing seed Arabidopsis thaliana cDNA
 clone M1B12 5', mRNA sequence.
 ACCESSION BE520240
 VERSION BE520240.1 GI:9778242
 KEYWORDS EST.
 SOURCE thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 393)
 White, J. A., Todd, J., Newman, T., Focks, N., Gilke, T., Martinez de
 Alarduya, O., Jaworski, J. G., Ohlroge, J. and Benke, C.
 A new set of Arabidopsis expressed sequence tags from developing
 seeds. The metabolic pathway from carbohydrates to seed oil
 Plant Physiol. 124 (4), 1582-1594 (2000)
 JOURNAL 20567808
 MEDLINE
 COMMENT Contact: Benning, C
 Dept. of Biochemistry & Molecular Biology
 Michigan State University
 224 Biochemistry, Michigan State University, East Lansing, MI 48824
 , USA
 Tel: 517 355 1609
 Fax: 517 353 9334
 Email: benning@msu.edu
 Michigan State University DNA Sequencing Facility Arabidopsis
 Biological Resource Center, The Ohio State University, 309 Botany &
 Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX:
 6142920603 TEL: 6142929371.
 FEATURES
 source
 1..393
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="M1B12"
 /clone_lib="Arabidopsis developing seed"
 /tissue_type="seed"
 /dev_stage="5-13 days after flowering"
 /lab_host="E.coli"
 /note="Organ: Developing seed; Vector: pBluescript SK-;
 Site 1: EcoRI; Site 2: XhoI"
 BASE COUNT 86 a 100 c 103 g 102 t 2 others

ORIGIN

Alignment Scores:

```
Pred. NO.: 1.4e-57
Score: 641.00
Percent Similarity: 97.71%
Best Local Similarity: 96.95%
Query Match: 27.85%
DB: 10
```

US-09-777-566A-2 (1-440) X BE520240 (1-393)

QY	107	ILLELLAASPVALLASPIGLARITHTRYGLYSTIRNGIYLALAPHELLAIIAGLYLEU	126
Db	393	ATTATTGCTGATGTCGAGCAGCCGTRACCCTAAMNACGGCGAAGCTTTGCCCCGGGNGT	334
QY	127	ALAPROASPYSALAIETHTVRVALNHTHGINALAASTHVSERSEPROASPROLEU	146
Db	333	GCACCTGACTGTGCATATACCGGTACATACCAGCGAGATAGCTCCAGTCCGATCCGTTA	274
QY	147	PHASNPRTLEULYTHRGILVYALCYSGINLEUASPANALAASTVALTHASPALAILE	166
Db	273	TTTATATCCTCTAAAACACTGCGGCTTTGCCAACATGGATTAACGCCAAGCTGACTAGCCGATC	214
QY	167	LEUSERIRGALAGLYSERTLEALLASPTHEHTHGINHISARGGLNTHRLAHPLEATG	186
Db	213	CTCAGCAGGGGAGAGGGGTCAATTCCTACTTACCGGGGCATCGGCATAACGGCGTTTCCG	154
QY	187	GLIUEUGLITRGVALLEUASNPHERGINSERANLEUCYLEULYSARGGLIUYSGIN	206
Db	153	GAATGTGAGACGGGTGCTTAATTTTCCGCAATCAACTGTGTGCTTAAACGTAGAAACAG	94
QY	207	ASPIGLUSERYSSERLEUTHRGINALALEUPROSERGIULEULYVALSERALLASPAEN	226
Db	93	GACCAAGAGCTGTTCATTATACGACAGCATTACATGGAAGTGAAGTGAGCCCGACAT	34
QY	227	VALSERLEUTHRGILYALAYALSERLEULAAASER	237
Db	33	GTCTCATTTAAACGCTGCGTAAAGCTCCATCA	1

RESULT 5	AA545747/c	LOCUS.	DEFINITION
AA545747	354 bp	mRNA	Linear
HBM5F1B4-REV	Human Bone Marrow Stromal Fibroblast	Homo sapiens	CDNA
clone HBM5F1B4 5',	mRNA sequence.		

FEATURES	Location/Qualifiers
source	1. .354

```

/sex="Male and Female"
/tissue_type="bone marrow"
/cell_type="stromal fibroblast"
/dev_stage="mixed"
/lab_host="X1-Blue"
/note="Vector: Bluescript; Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      77 a      95 c      95 g      95 t      others
ORIGIN

```

Alignment Scores:

```
Pred. No.: 5.71e-4
Score: 511.50
Percent Similarity: 94.968
Best Local Similarity: 94.128
Query Match: 22.228
DB: 9
```

US-09-777-566A-2 (1-440) x AA545747 (1-354)

QY	166	lleuenseeratalaglyglyserleuallaspheethinglyhisargglthr- ¹ alaph	185
	:::		
Db	352	gtcttcacagcagggcagagctca-atttctga-tttaccgggcattcgccmaacggcgctt	295
QY	185	earggluleuengluarvalleuasnpheproglinsersaneucysleuylsarggluly	205
Db	294	tcgggaactggmaacggggtgt-aattttcccaacaaactgttccttaaacctgagaa	236
QY	205	sglaspgluserserleuthnglnalileuprosergluleuylsvalseralaa	225
Db	235	acagagacgaacgtgtttatttaacgcagacattaccatcggaactcmaagtgtagggccga	176
QY	225	pasnvalserleuthrglyalavalserleuualsermetleuthrgluliepheule	245
Db	175	caagtntcattaaacccgtggcggaagcctgcacatcaatgcctgacggagatntttntcct	116
QY	245	uglnglnalaglnglymetprogluprcglityrpglytrglliehparsernaglntr	265
Db	115	gcacacagacacagggaatgccggagccgggggggaagaaagatcacccgattcacacacagtg	56
QY	265	pasntrhleuenseerleuuhisasnalaaglnpheyrleuenglinaqtthpro	283
Db	55	gaacacacctgtgaagttgcatatcaacggcccaatttatattgtgcacaaacgcaccaca	1

RESULT 6	BE436403/c	LOCUS	DEFINITION
	BE436403	320 bp	linear
	EST407481	tomato breaker fruit,	EST 18-MAY-2001
	clone c16c32c20,	mRNA sequence.	TIGR Lycopersicon
			esculentum cDNA

```

FEATURES      Location/Qualifiers
source        1. 320
              /organism="Lycopersicon esculentum"

```


100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 This clone is available through the Clemson University Genomics Institute
 Seq primer: T3.

FEATURES

source

location/Qualifiers
 1..208
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CDE61013"
 /clone_1lb="tomato breaker fruit"
 /tissue_type="pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptKmcUadapt; Site_1: EcoRI; Site_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research. Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 43 a 56 c 62 g 47 t
 ORIGIN

Alignment Scores:
 Pred. No.: 8.37e-28 Length: 208
 Score: 355.00 Matches: 68
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 15.42% Indels: 0
 DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BM412806 (1-208)

QY 82 HistyGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIstysGlyCysProGln 101
 |||||
 Db 206 CATTACCAACGCCAGCTCTGCTAGCCGACGATTGCTGCCGAAAAAGGCTGCCCCAG 147
 QY 102 SerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgIsthrGlyGluAla 121
 |||||
 Db 146 TCTGTCAGAGTCGCGATTATTGCTGATGTCGACGAGGTACCCGTAAACAGGCGAAGCC 87
 QY 122 PheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSer 141
 |||||
 Db 86 TTGCGCGCGCGGTGCGACCTGACTGTGCAATACCGTACATCCCGACGAGATACCTCC 27
 QY 142 SerProAspProLeuPheAsnPro 149
 |||||
 Db 26 AGTCCCGATCCGTATTATTATCTCT 3

RESULT 9
 BJO40796 546 bp mRNA linear EST 06-DEC-2001
 LOCUS BJO40796 NIBB Mochii normalized Xenopus neurula library xenopus
 DEFINITION laevis cDNA clone XLO47108 5', mRNA sequence.

ACCESSION BJO40796 GI:17388187
 VERSION BJO40796.1
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidea; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 546)
 Kitayama, A., Terasaka, C., Mochii, M., Ueno, N., Shin-I, T. and Kohara, Y.

TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan

Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tsuhin@genes.nig.ac.jp.
 location/Qualifiers
 1..546
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XLO47108"
 /clone_1lb="NIBB Mochii normalized Xenopus neurula library"
 /tissue_type="whole embryo"
 /dev_stage="stage 15"
 /note="Vector: pBSRN3; Site_1: NotI; Site_2: EcoRI; CDNAs were oligo-dT primed and directionally cloned. Staging according to Mueenkoop and Faber. Library is substracted and was constructed by N. Garrett and A.M. Zorn, (Wellcome/CRC Institute)."

FEATURES

source

BASE COUNT 116 a 145 c 142 g 138 t 5 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1.92e-25 Length: 546
 Score: 339.00 Matches: 66
 Percent Similarity: 98.51% Conservative: 0
 Best Local Similarity: 98.51% Mismatches: 1
 Query Match: 14.73% Indels: 0
 DB: 13 Gaps: 0

US-09-777-566a-2 (1-440) x BJO40796 (1-546)

QY 138 AlaAspThrSerSerProAspProLeuPheAsnProLeuYsthrGlyValCysGlnLeu 157
 |||||
 Db 345 GCGGTGACGTCACAGTCCGATCCGATATTAATCCCTAAACCTGGCTTGCAACTG 404
 QY 158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPhe 177
 |||||
 Db 405 GATACCGCAACGATGATGACGCAATCTCAGCAGGCGAGGCTAATTCGACTTT 464
 QY 178 ThrGlyHisArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnPheProGlnSer 197
 |||||
 Db 465 ACCGGGATGGGCAACGGCGTTTCCGAACTGGACGGGTATTTTCCGCAATCA 524
 QY 198 AsnLeuYsthrLeuYsthrGlu 204
 |||||
 Db 525 AACTGTGCTTAACGCTGAG 545

RESULT 10
 BG457132 657 bp mRNA linear EST 19-MAR-2001
 LOCUS BG457132 Phosphate starved leaf Medicago truncatula cDNA
 DEFINITION clone NF10C05PL 5', mRNA sequence.

ACCESSION BG457132
 VERSION BG457132.1 GI:13380457
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula

REFERENCE
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

1 (bases 1 to 657)
 Liu, J., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W., May, G.D. and Harrison, M.J.
 Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula phosphate-starved leaf library
 Unpublished (2000)

TITLE Unpublished (2000)
 JOURNAL Contact: Harrison MJ
 COMMENT Plant Biology Division
 The Samuel Roberts Noble Foundation
 2510 Sam Noble Parkway, Ardmore, OK 73402, USA
 Tel: 580 221 7325
 Fax: 580 221 7380
 Email: mjharrison@noble.org

Insert length: 657 Std Error: 0.00
 Plate: 100 row: C column: 05
 Seq primer: TCACACAGAAACGCTATGAC.

FEATURES
 SOURCE
 Location/Qualifiers
 1..657

/organism="Medicago truncatula"
 /db_xref="taxon:3880"
 /clone="NF100C05PL"
 /clone_lib="phosphate starved leaf"
 /tissue_type="leaf"
 /dev_stage="trifoliolate"
 /note="Vector: Lambda Zap; At the trifoliolate stage, M. truncatula plants were transplanted to phosphate-free sand and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20mM potassium phosphate. RNA was prepared from above ground tissues."

BASE COUNT
 ORIGIN
 187 a 172 c 138 g 160 t

Alignment Scores:
 Pred. No.: 8.05e-20 Length: 657
 Score: 287.00 Matches: 55
 Percent Similarity: 94.92% Conservative: 1
 Best Local Similarity: 93.22% Mismatches: 2
 Query Match: 12.47% Indels: 1
 DB: 12 Gaps: 0

US-09-777-566a-2 (1-440) x BG457132 (1-657)

OY 251 MetProgluProgluYTPGlyARtglethraspSerHisGIntPrAsnThrleuSer 270
 ::::|||||
 DB 481 ATTCGGAGCGGGGTGGGAAGATCACCATTACACCACTGGAACCTGTGTAAGT 540
 OY 271 LeuHisAsnAlaGlnPheTyrleuEnglnArgrThrProgluValAla-ArgSerArgAl 290
 :|||||
 DB 541 TTGCATACGCGCATTTTATTTGCTACACGACGCGAGGTTGCCCGGAGCGCGC 600
 OY 290 aThrProleuLeuSplcullleMetAlaAlaLeuThrProHisProProgluLys 308
 :|||||
 DB 601 CACCCCGTTATTATGATTCACAGACGCGTTCGCCGCCATCCACCGCAAAA 655

RESULT 11
 BE461872/c 179 bp mRNA linear EST 18-MAY-2001
 LOCUS EST413291 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
 DEFINITION BE461872
 ACCESSION BE461872
 VERSION BE461872.1 GI:9506174
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.
 1 (bases 1 to 179)
 AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Romming,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,D.J. and Tanksley,S.D.
 TITLE Generation of ESTs from tomato fruit tissue, breaker stage
 JOURNAL Unpublished (2000)
 COMMENT Contact: CUGI
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Email: <http://www.genome.clemson.edu/orders/index.html>
 5 prime sequence.

FEATURES
 SOURCE
 Location/Qualifiers
 1..179
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"

/db_xref="taxon:4081"
 /clone="CLEG40622"
 /clone_lib="tomato breaker fruit, TIGR"
 /tissue_type="Pericarp"
 /dev_stage="breaker"
 /lab_host="SOLR"
 /note="Vector: pBluescriptSKmCunadpt; Site 1: EcoRI; Site 2: XhoI; Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT
 ORIGIN
 40 a 47 c 51 g 41 t

Alignment Scores:
 Pred. No.: 8.99e-20 Length: 179
 Score: 278.00 Matches: 55
 Percent Similarity: 93.22% Conservative: 0
 Best Local Similarity: 93.22% Mismatches: 4
 Query Match: 12.08% Indels: 0
 DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x BE461872 (1-179)

OY 134 ValHisThrGlnAlaAspThrSerSerProAspProleuPheAsnProleuTyrThrGly 153
 :|||||
 DB 177 GTACATACCCAGCGAGATGCTCCGATCCGATTTATTCCTCTAAAACTGCG 118
 OY 154 ValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySer 173
 :|||||
 DB 117 GTTTCGACGCTGTGATACGTGAACGATCAGCGGATCTCAGACGAGGAGGTCA 58
 OY 174 IleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGlnArgValLeu 192
 :|||||
 DB 57 ATTGCTCACTTCACCGGCGATCGCAAAAGGGTTTCGCACTGGAACGGGTGCTT 1

RESULT 12
 BJ030203 508 bp mRNA linear EST 05-DEC-2001
 LOCUS BJ030203 NIBB Mochli normalized Xenopus neurula library Xenopus
 DEFINITION BJ030203
 ACCESSION BJ030203
 VERSION BJ030203.1 GI:17371629
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae; Xenopodinae; Xenopus.
 1 (bases 1 to 508)
 AUTHORS Kitayama,A., Terasaka,C., Mochli,M., Ueno,N., Shin-I,T. and Kohara,Y.
 TITLE Expressed genes in X. laevis embryo
 JOURNAL Unpublished (2001)
 COMMENT Contact: Tadasi Shin-I
 Center For Genetic Resource Information
 National Institute of Genetics
 1111 Yata, Mishima, Shizuoka 411-8540, Japan
 Tel: 81-559-81-6856
 Fax: 81-559-81-6855
 Email: tshin@genes.nig.ac.jp.

FEATURES
 SOURCE
 Location/Qualifiers
 1..508

/organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone="XL003h05"
 /clone_lib="NIBB Mochli normalized Xenopus neurula library"
 /tissue_type="whole embryo"
 /dev_stage="stage 15"
 /note="Vector: pBSRN3; Site 1: NotI; Site 2: EcoRI; cDNAs were oligo-dT primed and directionally cloned. Staging according to Nieuwkoop and Faber. Library is subtracted

Zoology Bldg., 1735 Neil Avenue, Columbus, OH 43210 USA, FAX: 6142920603 TEL: 6142929371.

FEATURES

source

```

/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="M1B12"
/clone_1fb="Arabidopsis developing seed"
/tissue_type="seed"
/dev_stage="5-13 days after flowering"
/lab_host="E. coli"
/note="Organ: Developing seed; Vector: pBluescript SK-,
Site_1: EcoRI; Site_2: XhoI"
BASE COUNT      70 a      57 c      63 g      64 t
ORIGIN

```

Alignment Scores:

Pred. No.:	6.26e-14	length:	25
Score:	225.00	Matches:	47
Percent Similarity:	92.31%	Conservative:	1
Best Local Similarity:	90.38%	Mismatches:	4
Query Match:	9.77%	Indels:	0
DB:	10	Gaps:	0

US-09-777-566A-2 (1-440) x BE520241 (1-254)

QY 1 MetLysAlaIleLeuLeuProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 96 ATAAAGGCATCTTAATCCATTTTAACTCTTCGATTCGGTAACCCCAATCTGCA 155
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 156 TTGCTCAAGATGAGCGGAGGTGGAAGCTGGAAATGGTGGATTCGACGTGCATGGT 215
QY 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGln 52
Db 216 GTGCGTGTCTCCACCAAGGACAGGACACTGATGACAG 251

RESULT 15

AU004475

LOCUS	696 bp	mRNA	linear	EST 19-JAN-1999
AU004475				
DEFINITION	Bombyx mori p50(Daizo)	Bombyx mori	cDNA clone	ws20511,

ACCESSION

VERSION AU004475.1 GI:4161846

SOURCE

ORGANISM *Bombyx mori*

Bombycoidea; Bombycidae; Bombyx.

REFERENCE 1 (bases 1 to 696)

AUTHORS Mita, K., Morimyo, M., Shimada, T., Okano, K. and Maeda, S.

TITLE	Establishment of cDNA database c
FORM	1000

JOURNAL
Unpublished (1999)

COMMENT: Contact: Mita K
Gordon Research Center

Genome Research Group
National Institute of Diabetes and Digestive and Kidney Diseases

National Institute of Radiologic
Address 4-0-1 Tama-shi Chiba 263-

Angawa 4-9-1, Inage, Chiba 263-
Email: km4a@n4s.cc.jp

EMAIL: kmlaenlis.go.jp
PROJECT = 'CPST project by TEM'

PRODEC1 - CRESI project by user
Location/Qualifiers

Source	Location/Qualifiers
Exhibitions	1 696

Source: 1 : 050
/organism="Bombyx mori"

```
/organismi - bonnyz inoia
/strajne="50(Dajzo)"
```

```

/serial=psv(bazco,
/db xref="taxon:7091"

```

```
/usr/bin/canon /opt  
/clone="ws20511"
```

/clone 11b="Bombvix mort

BASE COUNT	237 a	132 c	136 q

ORIGIN

Alignment Scores:

Pred. No.:	1.11e-07	length:	696
Score:	172.50	Matches:	49
Percent Similarity:	44.69%	Conservative:	31
Best Local Similarity:	27.37%	Mismatches:	78
Query Match:	7.49%	Indels:	21
DB:	9	Gaps:	5

US-09-777-566A-2 (1-440) x AU004475 (1-696)

QY	28	LeuLysLeuLysLeuSerValLeuValLeuValSerValGlnSerGlyValAspValAspValProThrLeuGlyAla	47
Db	130	CTACGGCTTGAGCAAGACTATTATTATTTTAACTGCTGCACAACTTAAGAGTTGCA-----CTA	183
QY	48	ThrGlnLeuMetGlnAspValThrProAspAlaTrpProThrTrpProValLysLeuGly	67
Db	184	TCMAAACAAATAGATCAATATACAAACGAAACAAATATTATTGCGAAATGGTCCCAAGACCCGGC	243
QY	68	TrpLeuThrProArgLys-----GlyLysLeuIleAlaTyrLeuGlnLysTyr	83
Db	244	ATGCTGACGAAAAAAGAGACCTTATTGGAAGGTCAATGACGAGATGAGCAATTAACGAAATGG	300
QY	84	GlnArgGlnLeuValAlaAspGlyLeuLeuAlaLysGlyCysArgProGlnSerGly	103
Db	304	ATGAGAGAGAAATCAACTTTTACCSCGGAACG-----TGTCCGAGACAAAGAA	348
QY	104	GlnValAlaIleIleLeuAspValAspGlnArgTrpArgLysThrGlyAlaIleAla	123
Db	349	ACCGTTTGATGTTTATCGCATTAATAAGACACCAACCAATGACTGCGAAGGCAATTGGTC	408
QY	124	AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerPro	143
Db	409	GATGCCGCAATTCCTGACTGCAATGTAATTCGTGAAAGCATTAAGAAAGATTTTGATATACAT	466
QY	144	AspProLeuPheAsnProLeuLeuYthrGlnGlyValCysGlnLeuAspAsnAlaAsnValThr	163
Db	469	GACATCATCTTCATATCTGCT-----ATTCAATACACCAAC	504
QY	164	AspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlnLysArgGln--	182
Db	505	GAAAGCTACAAACAAAAGTAAGTCTCGAAGAAATCGAGAAATGCTGCCCAATTTGTAACATA	564
QY	183	---ThrAlaPheArgGlyLeuGlnLysAlaValLeuAsnPheProGlnSerAsnLys	200
Db	565	ACAGATGCTCGACAAATATTAGACAGATTTATCGCATTAAGAAACCTCTAAATATTTGCG	621

Search completed: June 12, 2003, 13:58:39
Job time : 1938 secs

Alignment Scores:

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 10:23:20 ; Search time 70 Seconds
(Without alignments)
1927.683 Million cell updates/sec

Title: US-09-777-566a-2
Perfect score: 2302
Sequence: 1 MKAILPFLSLIPRPOSA.....IVNEARIPACSLRHHHHH 440

Scoring table: BLOSUM62
Xgapop 10.0, Ygapext 0.5
Xgapop 6.0, Ygapext 7.0
Delop 6.0, Delext 7.0

Searched: 441362 seqs, 153338381 residues
Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+ p2n model -DEV-xlh
-O/cg2_1/USPTO.spool/US0977566/rnat_05062003_154804_6427/app.query.fasta_1.583
-DB-Issued_Patents_NA -OPMT-fastap -SOFTX-rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS-bits -START=1 -END=1 -MATRIX-blosum62 -TRANS-human40.cdi
-LIST=45 -DOCALIGN=200 -THR-SCORE-pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE-LOCAL -OUTPMT-plo -NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US0977566.ecgn.1.1.40.tunat.05062003_154804_6427 -NCPU=6 -ICPU=3
-NO_MMP -LARGEQUERY -NEG_SCORES=0 -NAT -DSPBLOCK=100 -LONGLOC
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_NA:*
1: /cg2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cg2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cg2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cg2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cg2_6/ptodata/1/ina/PCTUS.COMB.seq:*
6: /cg2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2302	100.0	1323	3	US-09-259-214-1
2	2302	100.0	1323	4	US-09-318-528-1
3	2302	100.0	1323	4	US-09-291-931-1
4	2194.5	95.3	1272	2	US-08-910-798-1
5	352	15.3	5975	1	US-08-920-812-23
6	352	15.3	5975	1	US-08-920-827-23
7	352	15.3	5975	1	US-08-921-177-23
8	352	15.3	5975	1	US-08-362-57C-23
9	352	15.3	5975	2	US-08-920-828-23
10	119	5.2	1571	4	US-08-868-435-32
11	119	5.2	1571	4	US-08-744-231-32
12	119	5.2	1571	4	US-09-044-718-10

13	117	5.1	1588	2	US-09-146-283-1	Sequence 1, Appl
14	117	5.1	1588	3	US-08-579-823A-1	Sequence 1, Appl
15	117	5.1	1588	4	US-09-344-195-1	Sequence 1, Appl
16	117	5.1	3061	2	US-08-692-787-47	Sequence 47, Appl
17	117	5.1	3061	4	US-09-097-199-47	Sequence 47, Appl
18	112	4.9	11604	4	US-09-385-028-13	Sequence 13, Appl
19	112	4.8	15079	3	US-09-385-028-1	Sequence 1, Appl
20	111.5	4.8	80161	3	US-09-036-987A-1	Sequence 1, Appl
21	111.5	4.8	80161	3	US-09-370-700-1	Sequence 1, Appl
22	108	4.7	5153	5	PCT-US95-04910-8	Sequence 8, Appl
23	106.5	4.6	3147	4	US-08-887-534A-84	Sequence 8, Appl
24	105	4.6	1912	4	US-08-868-435-11	Sequence 11, Appl
25	105	4.6	1912	4	US-08-744-231-11	Sequence 11, Appl
26	104	4.5	1642	3	US-08-993-359-27	Sequence 27, Appl
27	103.5	4.5	2071	1	US-07-923-724-1	Sequence 1, Appl
28	103.5	4.5	2071	2	US-08-609-426A-1	Sequence 1, Appl
29	103.5	4.5	2071	2	US-08-374-652C-3	Sequence 3, Appl
30	100	4.3	12508	4	US-09-655-270A-1	Sequence 1, Appl
31	100	4.3	12523	4	US-09-651-941-1	Sequence 1, Appl
32	100	4.3	12523	4	US-09-955-597-1	Sequence 1, Appl
33	100	4.3	15378	3	US-08-785-420-1	Sequence 1, Appl
34	99.5	4.3	1845	4	US-08-868-435-30	Sequence 30, Appl
35	99.5	4.3	1845	4	US-08-744-231-30	Sequence 30, Appl
36	99.5	4.3	1845	4	US-09-044-718-7	Sequence 7, Appl
37	99	4.3	3836	2	US-08-216-260-1	Sequence 6, Appl
38	99	4.3	28958	1	US-08-258-261B-6	Sequence 6, Appl
39	99	4.3	28958	1	US-08-456-837-6	Sequence 6, Appl
40	99	4.3	28958	1	US-08-457-342-6	Sequence 6, Appl
41	99	4.3	28958	1	US-08-457-646A-6	Sequence 6, Appl
42	99	4.3	28958	1	US-08-458-076A-6	Sequence 6, Appl
43	99	4.3	28958	1	US-08-764-233A-4	Sequence 4, Appl
44	99	4.3	28958	1	US-08-457-335A-6	Sequence 6, Appl
45	99	4.3	28958	1	US-08-729-214-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-09-259-214-1
Sequence 1, Application US/09259214A
Patent No. 610719
GENERAL INFORMATION:
APPLICANT: Kretz, Keith
TITLE OF INVENTION: NOVEL PHYTASE
FILE REFERENCE: DIVER370-1
CURRENT APPLICATION NUMBER: US/09/259, 214A
CURRENT FILING DATE: 1999-03-01
EARLIER APPLICATION NUMBER: 08/910, 798
EARLIER FILING DATE: 1997-08-13
NUMBER OF SEQ ID NOS: 4
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1
LENGTH: 1323
TYPE: DNA
ORGANISM: Escherichia coli
FEATURE:
NAME/KEY: CDS
LOCATION: (1)...(1320)
NAME/KEY: misc_feature
LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-09-259-214-1
Alignment Scores:
Pred. No.: 3.03e-239
Score: 2302.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 100.00%
DB: 3
Gaps: 0
US-09-777-566a-2 (1-440) x US-09-259-214-1 (1-1323)

```

OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGCATCTTAATCCATTTTATCTCTTCGATTCGCTTAACCCCGCAATCTGCA 60
OY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTACCCGAGAGCTGAAGCTGGAAAGTGGTGATGTGATGCTATGATGCT 120
OY 41 ValAlaGlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCCACCAAGCAAGCCAGCACTGATGAGATGTCAGATCCCGAGAGCATGGCCA 180
OY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCGTGGCCGGTAAACTGGTGGCTGACACCGCGNGGTGGAGCTAATGCGCCATCTTC 240
OY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysPro 100
Db 241 GGACATTTCCAAACGCGCAGCTCTGTGTTACCGAGGATTTGCTGGCAAAAGGGCTGCCG 300
OY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGTGTCAGATCGCATTTATTCGTGATGTCAGAGCGATACCCGTAACAGAGCGCA 360
OY 121 AlaPheAlaIaGlyLeuAlaProAspGlyAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCGTTCGCGCGCGGCTGGACCTGACTGCTCAATACCCGTAACACCGAGCATGACG 420
OY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCACATCCCGATCCCTTTATTTAATCCTTAAACATGGCGTTTGGCAACTGATACCGG 480
OY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACGCGATCTCAGAGCGGAGAGGGTCAATGCTGACTTACCGCGCAT 540
OY 181 ArgGlnThrAlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuLys 200
Db 541 CGGCAACAGGGGTTTCCGAACCTGGAAAGCGGTGCTTAATTTCCCAATCAAACTTGGC 600
OY 201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlnLeu 220
Db 601 CTTAAAGCTGAAGAAACAGAGCAAGAGCTGTCATTAAAGCGAGCTTACCATCGGAATC 660
OY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGGTGACCGCGCAATGTCTCATTAACCGGTGCGGTGATGCTGATCAATGCTGACG 720
OY 241 GluIlePheLeuLeuGlnGlnAlaGlnIleGlyMetProGlnProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAAAGCAAGAGGAAATGCCGAGCGGGGTGGGAAGGATCAC 780
OY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuLeuGln 280
Db 781 GATTTCACACGAGTGAACACCTTGCTAGTTTGGCTAATACGGGCATTTTATTTGCTACA 840
OY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCAGAGGTGGCCGAGCGCGCACCCCGTATTGATGATCATGACATGACAGCG 900
OY 301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCGCCATCCACCCCAAAACAGCGGATGGTGAGCATTAACCCACTTCATCTACTG 960
OY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaIleGlnIleLeuAsnTrp 340
Db 961 TTTTATTTGCGGAGACAGATACATCTGCGCAATCTCGGCGCGCATCTGAGCTCAACTGG 1020
OY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGlnArgTrp 360
Db 1021 ACGGTTCCTCCGCTCAGCCGATACACGCGCGAGGTGTGAACGTGTTTGAACGCTGG 1080
OY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

```

```

Db 1081 CGTCGCTAAGCAAGCAACAGCCAGTGAATCAGTTTGGCTGCTTCAGACTTTACAG 1140
OY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1141 CAGATGCGGTGATTAACCCCGCTGCTCATTAATTAATACCCCGCGAGAGTGAACCTGACC 1200
OY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnIleGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGAGATGGAAGAGCAAGAAATGCGCAGGCGATGTGTGTTGGCAGGTTTACGCA 1260
OY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db 1261 ATCGTAATGAAGCAGCATACCGCGGTGAGATTGAGATCTCATCACCATCACCATCAC 1320

RESULT 2
US-09-318-528-1
; Sequence 1, Application US/09318528
; Patent No. 6183740
; GENERAL INFORMATION:
; APPLICANT: Kretz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/318,528
; CURRENT FILING DATE: 1999-05-25
; EARLIER APPLICATION NUMBER: 09/291,931
; EARLIER FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-318-528-1

Alignment Scores:
Pred. No.: 3,03e-239 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-777-566A-2 (1-440) x US-09-318-528-1 (1-1323)
OY 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGAGCATCTTAATCCATTTTATCTCTTCGATTCGCTTAACCCCGCAATCTGCA 60
OY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValIleValSerArgHisGly 40
Db 61 TTCGCTCAGAGTACCCGAGAGCTGAAGCTGGAAAGTGGTGATGTGATGCTATGATGCT 120
OY 41 ValAlaGlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 121 GTGCGTGTCCACCAAGCAAGCCAGCACTGATGAGATGTCAGATCCCGAGAGCATGGCCA 180
OY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
Db 181 ACCGTGGCCGGTAAACTGGTGGCTGACACCGCGNGGTGGAGCTAATGCGCCATCTTC 240
OY 81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyLysPro 100

```

```

Db      241 GGACATTAACAAGCCAGCGTGTGGTACCGACGATTGTCGGCAAAAAGCGCTGCCG 300
Qy      101 GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db      301 CAGCTGTGTCAGTCCGCGATTATTCGTCTATCTCGACGAGCGCTACCCGTAACAGGCCAA 360
Qy      121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      361 GCCTTCGCCCGCGGGCTGGCACCTGACTGTCATAAACGTAACATACCAAGCAGATACG 420
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      421 TCCAGTCCCGGATCGGTATTTAAATCTCTAAAACTGGCGTTTGCAACTGGATAACCG 480
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      481 AACGTGACTGACGGGATCTCAGCAGCAGGAGGTAAATTCGCTTACCCGGCAT 540
Qy      181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db      541 CGGCAAAACGGCGTTCCGCAACTGGAAACGGGTCTTAATTTCCGCAATCAAACTTGGC 600
Qy      201 LeuLysArgGluLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      601 CTTAAAGCTGAGAAACAGACGAAAGCTGTTCAATTAAAGCAGGCAATTCGGAACCTC 660
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db      661 AAGGTGACCGCCGCAAAATGCTCATTAACCGGTGGGTAACCTTCGCTCAATGCTGACG 720
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnLysMetProGluProGlyLysArgIleThr 260
Db      721 GAGATATTTCTCTGCAACAGACAGAGGATGCCGAGCGGGGTGGGGAAGATCAAC 780
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGln 280
Db      781 GATTCAACACAGTGGAAACCTTCTAAGTTGCAATAACGGCAATTTATTTGCTACAA 840
Qy      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIle 300
Db      841 CGCACCGCAGAGTGGCCCGCAGCCGCGCACCCCGTTATGATTTATCATGTCACG 900
Qy      301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db      901 TTGACGCCCATTCACCCGCAAAACAGCGCATGGTGTGACATTAACCACTTCAGTACTG 960
Qy      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrp 340
Db      961 TTTATTTGCCGGACAGATACTAATCTGGCAATCTCGCGCGGCGCACTGGAGCTCAACTGG 1020
Qy      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1021 ACGGTCCCGCTCACCGCGATTAACAGCCGCGAGGTGTGAATGCTTTTGAAGCGTGG 1080
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CGTGGCGTAACGATTAACAGCAGTGTGAGTTGCTGCTGCTTCAGACTTTTACAG 1140
Qy      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db      1141 CAGATGCGGTGTAACACCGCTCTGATTAATACGCGCGCGGAGAGTGAACACTGGACC 1200
Qy      401 LeuAlaGlyCysGlnGluArgAsnAlaGlnLysMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGGCAGGATGTGAAGACGAAATGCGAGGCGATGTGTGTTGGCAGGTTTACGCA 1260
Qy      421 IleValAsnGlnAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db      1261 ATCGTGAATGAAGACAGCATACCGGCGTGTGAGTTGAATCTCAACATCACCATCAC 1320

```

RESULT 3
 US-09-291-931-1
 ; Sequence 1, Application US/09291931A

```

; Patent No. 6190897
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Keith
; TITLE OF INVENTION: NOVEL PHYTASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/09/291,931A
; CURRENT FILING DATE: 1999-04-13
; EARLIER APPLICATION NUMBER: 08/910,798
; EARLIER FILING DATE: 1997-08-13
; EARLIER APPLICATION NUMBER: 09/259,214
; EARLIER FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature
; LOCATION: (1)...(1323)
; OTHER INFORMATION: n = A,T,C or G
US-09-291-931-1

Alignment Scores:
Pred. No.: 3,03e-239 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 4 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-291-931-1 (1-1323)
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      1 ATGAAGCGAGCTTATATCCATTATTTATCTCTTGATTCGTTAACCCCGCAATCTGCA 60
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValAlaIleValSerArgHisGly 40
Db      61 TTGCTGACAGTGAACCGGAGCTGAAGCTGAAAGTGTGTGATTTGTGATGCTATGCT 120
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db      121 GTGCTGCTCCAAACCAAGGCCACGCACTGATGAGATGTACCCCAAGCCATGGCCA 180
Qy      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db      181 ACGTGGCGGTAAACTGGGTGGCTGACACCGCGNGGTGGTGAAGCTTATGCGCTATCTC 240
Qy      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlySerPro 100
Db      241 GGACATTTACCAACGGCACCGTGTGTACCGCAGGATTTGTGGCGAAAGGCGTGGCCG 300
Qy      101 GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db      301 CAGTGTGTCAGTCCGCGATTATTCGTCTATCTCGACGAGCGGTACCCGTAACAGGCCAA 360
Qy      121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      361 GCCTTCGCCCGCGGGCTGGCACCTGACTGTCATAAACGTAACATACCAAGCAGATACG 420
Qy      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db      421 TCCAGTCCCGGATCGGTATTTAAATCTCTAAAACTGGCGTTTGCAACTGGATAACCG 480
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      481 AACGTGACTGACGGGATCTCAGCAGCAGGAGGTAAATTCGCTTACCCGGCAT 540
Qy      181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200

```

```

Db      541 CGGCAACGGCGTTTCGGCAACTGGAACTGGTCTTAATTTCCGCAATCAAACTTGTGC 600
Qy      201 LeuysargGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db      601 CTTAAAGCTGAGAAACAGACGAAAGCTGTTCATTACCGCAGCATTTACATCGGAATC 660
Qy      221 LysValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAlaSerMetLeuThr 240
Db      661 AAGGTAGCGCCGCAATGTCTCAATTAACCGGTGCGTAACCTCGCATCATCGTGCAGC 720
Qy      241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGluIleThrGlyArgIlePhe 260
Db      721 GAGATATTTCTCTGCAACAGACAGGGAATGCCGAGCCGGGGTGGGAAAGATCCACC 780
Qy      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db      781 GATTACACACGATGGAAACACTTGTAGTTGATTAACGGGCATTTTATTGTGTACAA 840
Qy      281 ArgThrProGluValAlaAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db      841 CGCAGCGCCAGAGGTGGCCCGCAGCCGCCACCCCGTTATTTGATGATGATGCGACAGC 900
Qy      301 LeuThrProHisProGlnProGlnIleValAlaTyrGlyValIleThrLeuProThrSerValLeu 320
Db      901 TTGACGCCCATCCACCGCAAAACAGCGCATATGATGATGATTAACCATTCAGTACTG 960
Qy      321 PheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db      961 TTTATTTGCCGACACGATTAATCTGTGCAATCTCGCGCGGCGACCTGAGCTCAACTCG 1020
Qy      341 ThrLeuProGlnIleProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db      1021 ACGTCTCCCGTCAGCCGGATACACAGCGCCAGGTGTGAATCTGTTGAACGCTGG 1080
Qy      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db      1081 CGTGGCGTAACGATACAGCAGCATGATTCAGTTGCTGCTGCTTCACATCTTACAG 1140
Qy      381 GlnMetArgAspLysTrpProLeuSerLeuAsnThrProProGlyGlyValIleLysLeuThr 400
Db      1141 CAGATGCGTGAATAAACGCCCTGTCAATTAATACGCCGCCGAGAGTGAATAACTGACC 1200
Qy      401 LeuAlaGlyCysGluArgAsnAlaGlnIleMetCysSerLeuAlaGlyPheThrGln 420
Db      1201 CTGCGAGGATGTGAAGACGAAATGCGAGGCGCATGTGCTGTGCGAGGTTTACGCAA 1260
Qy      421 IleValAsnGlnAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
Db      1261 ATCGTGAATGAAGACGATACCGCGGTGCAAGTTGAGATCTCATACCATCACCATCAC 1320

RESULT 4
US-08-910-798-1
: Sequence 1, Application US/08910798
: Patent No. 5876997
: GENERAL INFORMATION:
: APPLICANT: KRETEZ
: TITLE OF INVENTION: NOVEL PHYTASE
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Fish & Richardson P.C.
: STREET: 4225 Executive Square, Suite 1400
: CITY: La Jolla
: STATE: California
: COUNTRY: US
: ZIP: 92037
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/910,798
: FILING DATE: August 13, 1997

```

```

: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: HALL, PH.D., LISA A.
: REGISTRATION NUMBER: 38,347
: REFERENCE/DOCKET NUMBER: 09010/029001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 619/678-5090
: TELEFAX: 619/678-5070
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1272 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: IMMEDIATE SOURCE:
: CLONE: PHYTASE
: FEATURE:
: NAME/KEY:
: LOCATION:
: US-08-910-798-1

Alignment Scores:
Pred. No.: 1,199-227 Length: 1272
Score: 2194.50 Matches: 423
Percent Similarity: 96.14% Conservative: 0
Best Local Similarity: 96.14% Mismatches: 0
Query Match: 95.33% Indels: 17
DB: Gaps: 1

US-09-777-566a-2 (1-440) x US-08-910-798-1 (1-1272)
Qy      1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db      1 ATGAAAGGAGCTATATCCCAATTTTATCTCTTGATTCGCTTAAACCCCAATCTGCA 60
Qy      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValAlaIleValSerArgHisGly 40
Db      61 TTCGCTCAGAGTACCGCCGAGCTGAAGCTGGAAGTGTGGTATGTCATGTCATGGT 120
Qy      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAlaThrProAspAlaTrpPro 60
Db      121 GTGCGTGTCCACCAAGGCCAGCACTGATCGAGATGTCACCCAGACGATGGCCA 180
Qy      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db      181 ACTCGCCGGTAAAACTGGTTGCTGTACACCGCGNGTGTGACATATGCGCTATCTC 240
Qy      81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db      241 GGACATTACCAACGCGCAGCTGTGTACCCAGGATGTGGCGAAAAAGGCGTGGCCG 300
Qy      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgLysThrArgLysGlu 120
Db      301 CAGTGTGTGTCAGTGTGCGATATATTCATGTGCGAGCGCATACCCGTAACAGCGC--- 357
Qy      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db      358 -----CAAGCAGATACG 369
Qy      141 SerSerProAspProLeuPheAsnProLeuLysTrpGlyValCysGlnLeuAspAsnAla 160
Db      370 TCCAGTCCCGATCCGTTATTTAATCTCTTAAACAGCGTTTGGCAACTGGATTAACCGC 429
Qy      161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db      430 AACGTGATGACGCGATCTCTACGACGAGGAGAGGCTCAATTGCTGACTTTACCGGGAT 489
Qy      181 ArgGlnThrAlaPheArgGluLeuGlnArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db      490 CGGCAACGCGGTTTCCGGAACGTGAACGAGGCTGCTAATTTTCGCAATCAAACTGTGC 549
Qy      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220

```


|||||
Db 550 CTTAAACGTGAGAAACAGACGAAACCTGTCATTAAACGACGCAATACCATCGAATC 609
OY 221 LysValSerAlaAspAsnValSerLeuThrGlnValValSerLeuAlaSerMetLeuThr 240
Db 610 AAGGTGAGGGCCGACATATCTCATTAACGGGGGCGGTAAAGCTCCATCAATCAAGTGAGC 669
OY 241 GlnLeuPheLeuLeuGlnGlnAlaGlnGlyMetProGlnProGlyTyrPglArgGlyLeuThr 260
Db 670 GAGATATTTCTCCTCGAACAAGACAGCGGGAATCCCGGAGCCGGGGGCGGGAAGATCACC 729
OY 261 AspSerHisGlnTTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 730 GATTCACACAGAGGAAACCTGTCAATGTTTCATACGCGCAATTTATTTGTCATACAA 789
OY 281 ArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuMetAlaAla 300
Db 790 CGCAGCCGAGAGTTCCCGCAGCCGCGCCACCCGCTTATGATTTGATTCATGGCAGCG 849
OY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlnValThrLeuProThrSerValLeu 320
Db 850 TTGACGCCCATCCACCGCAAAAACAGGCGTATGCTGTCATTAACCTTACATGACTG 909
OY 321 PheLeuAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlnLeuAsnTyr 340
Db 910 TTTATTCGCCGACACAGTACTATCTGCAATCTCGGGCGGCGCTGAGGACTCACTG 969
OY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGlnArgTyr 360
Db 970 ACGCTTCCCGGTACGCGGATTAACAGCGCCGCGGCTGTAACCTGCTGTTTGAACCTCG 1029
OY 361 ArgArgLeuSerAspAsnSerGlnTrrPileGlnValSerLeuValPheGlnThrLeuGln 380
Db 1030 CCGTGGCTAAGGATTAACAGCCAGTGTGATTCAGGTTCCGTGCTTCCAGACTTTCACG 1089
OY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1090 CAGATCGTGATTAACGCGCGCTGTCATTAAATACCGCCGCGGAGGAGGAAACGACG 1149
OY 401 LeuAlaGlyCysGlnGlnArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1150 CTGGCGAGATGTGAAGAGCAATGGCAGGCGCATGTGTCGTGACAGGTTTTTACGCCAA 1209
OY 421 IleValAsnGlnAlaArgLeuProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1210 ATCGTGAATGAAGCAGCATACCGGCTGAGTTGAGATCTCATCACCATCCATCAC 1269
RESULT 5
US-08-920-812-23
; Sequence 23, Application US/08920812
; Patent No. 5763188
; GENERAL INFORMATION:
; APPLICANT: Ohno, Tsuneya
; APPLICANT: Matsuhisa, Akio
; APPLICANT: Uehara, Hirotsugu
; APPLICANT: Ede, Soji
; TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
; STREET: 6300 Sears Tower, 233 South Wacker Drive
; CITY: Chicago
; STATE: Illinois
; COUNTRY: United States of America
; ZIP: 60606-6402
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/920,812
; FILING DATE: 29-AUG-1997

CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
; TELEX: 25-3856
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella pneumoniae
; STRAIN: Clinical isolate KI-50
US-08-920-812-23
Alignment Scores:
Pred. No.: 2,366-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 1 Gaps: 14
US-09-777-566a-2 (1-440) x US-08-920-812-23 (1-5975)
OY 2 LysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrPro---GlnSerAla 20
Db 578 GAGGGGTGTACGCTGTTATTCCTGCGCGCTGCGCTGCTGCGGCGGTCAATATGCC 637
OY 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 638 GCCGCC-----GCGGACTGCGACGTGAGAAAGTGGTCAGCTCAGCCGCCACAGGT 668
OY 41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaTyr 59
Db 689 ATTCGTCGCCGACAGCGCGGCAACCGGGAAGCCATCGACGCCGCCACGCCGACCGTGG 748
OY 60 ProThrTrrProValLysLeuGlyTrrPleuThrProArgGlyGlyLeuIleAlaTyr 79
Db 749 ACCGAGTGCACACCCATGACGGGAGACTCACCAGGCATGCTATGCCGCGGTGTCAC 808
OY 80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
Db 809 AAGGGCGTGGAGAGCCGACATTAACCGCCAGCTCGCTGCTG---CAGGCCGAGTGC 865
OY 100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgLysThrGly 119
Db 866 CCACAGCGGAGTGCATATACGTGCGCGCCACGCCCTGCGACGCGACCGACGACCGCC 925
OY 120 GlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db 926 CAGGCGCTGTGATGAGCGCTTCCCGGCTGCGGGCTGCTATTCATTTATGTC----- 979
OY 140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValLysGlnLeuAspAsn 159
Db 980 AGCGGGGATGCCATCCCTGTTTCAGACCGACAGTTCGCCGCCACGCAACGCCGCC 1039
OY 160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlyLysSerIleAlaAspPheThrGly 179
Db 1040 GCCCGCCA-GCTGCGCGGTGAAGAGAGAGCGCGG-----GATTCGCC-GCA 1085
OY 180 HisArgGlnThrAlaPheArgGlyLeuGlnArgValLeuAsnPheProGlnSerAsnLeu 199
Db 1086 GGTGCGCAGGCG-----CTGGGCCGACCATTCAGCTATTGAACAGCGGCTT 1133


```

Db      1086 GGTCCGACGACG-----CTGCGCGCCACCATCATCATTTGAACAGCGCGTT 1133
               |||||
QY      200 CysLeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerglu 219
               |||||
Db      1134 TGT-----CAGCCCATATAGCCCTGCGCATCTGCATACC---CCGTGGCAG 1178
               |||||
QY      220 LeuLysValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuLysSerMetLeu 239
               |||||
Db      1179 GTCGACGACAGCAAAAGTGGGAAGACACCATTTAGCGGACGTGAGCTGATGGCCAATATG 1238
               |||||
QY      240 ThrGlnLeuPheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlyTyrGly 257
               |||||
Db      1239 GTGAGACGCTCGCTCGCTGAGGTGAAACCTGCTCTGACGACGAGTGGCGTGGGG 1298
               |||||
QY      258 ArgGlyLeuAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyr 277
               |||||
Db      1299 AAGATCACCCAGGACGAGATCACCGCCCTGCTGCTTAAAGGAAACTACGAT 1358
               |||||
QY      278 LeuLeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuLeu 297
               |||||
Db      1359 CTGAGTAAACGATGTGTATACCGGCAAAACGCGGTGCTGCTGCTCAACGCTATG 1418
               |||||
QY      298 MetAlaAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThr 317
               |||||
Db      1419 CTCGACGCGCTCAACCCGAGCGAATCGAACGT-----ACGCTGGCTGCT 1463
               |||||
QY      318 SerValLeuPheLeuAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlu 337
               |||||
Db      1464 GCT-----GTTGGCCATGATACCAATATGCTCATGTCGGGACGCGTGAAGAC 1511
               |||||
QY      338 LeuAsnTrpThrLeuProGly---GlnProAspAsnThrProProGlyGluLeuVal 356
               |||||
Db      1512 TTTACTGCGACACTGCGCGGTACACCGCGGGAATATCCCGCGGACGACCTGGTG 1571
               |||||
QY      357 PheGluArgTrpArgLeuSerAspAsnSergLysIleGlnValSerLeuValPhe 376
               |||||
Db      1572 CTGAGACGCTGCGACACCGAAGACGAGAGACGCTATCTCGGCTCATTTCCAGGCC 1631
               |||||
QY      377 GlnThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
               |||||
Db      1632 CAGGCGCTGACGACGCTGCTGCTGACGACCGCGGACCGCCGATGCTGCGT 1691
               |||||
QY      397 ValLysLeuThrLeuAlaGlyCysGlnGluIleArgAsnAlaGlnGlyMetCys----- 413
               |||||
Db      1692 CAGGAGTGGCATCACCGCGCTGCTGACGACCGATGCGTACGCTGTGCTCCCTTCAG 1751
               |||||
QY      414 ---SerLeuAlaGlyPheThrGlnIleValAsnGluAlaArgIleProAla 429
               |||||
Db      1752 GCGGCTATTACCGCCCTGCTGACGCTATCGACCGATCATCCGCCCGCGC 1802
               |||||

```

RESULT 7
US-08-921-177-23
Sequence 23, Application US/08921177
Patent No. 5798211

GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/921,177
; FILING DATE: 29-AUG-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/362,577
; FILING DATE: 27-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Rin-Laures, Li-Hsien
; REGISTRATION NUMBER: 33,547
; REFERENCE/DOCKET NUMBER: 19036/32420
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/474-6300
; TELEFAX: 312/474-0448
;
; INFORMATION FOR SEQ ID NO: 23:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Klebsiella pneumoniae
; STRAIN: Clinical Isolate KI-50
;
; US-08-921-177-23

```

```

Alignment Scores:
Pred. No.: 2,36e-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 1 Gaps: 14

```

```

US-09-777-566a-2 (1-440) x US-08-921-177-23 (1-5975)
QY      2 LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrPro---GlnSerAla 20
               |||||
Db      578 CAGGGCTGTACGCTGTATATGCTGCGCGCTGCTGCTGCGCTGCAATCTGCC 637
               |||||
QY      21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
               |||||
Db      638 GCCGCG-----CGGACGTGGCAGCTGGAGAAAGTGTGACGTCAKCGCAGGT 688
               |||||
QY      41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaTrp 59
               |||||
Db      689 ATTCGTCGCCGACGCGCGCAACCGGAAAGCATGAGAGCCGCCACCGCGGACCGTGG 748
               |||||
QY      60 ProThrTrpProValLysLeuGlyTyrPheLeuThrProArgGlyGlyLeuIleAlaTyr 79
               |||||
Db      749 ACCGAGTGGACCAACCCATGACGGGAGCTCACCGCCATGCTATCCCGCGTGCAC 808
               |||||
QY      80 LeuGlnHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
               |||||
Db      809 AAGGCGCTGGGAGAGCCAGCATTTACCGCCAGCTGGCTGCTG---CAGCGCGATGC 865
               |||||
QY      100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspValAspGlnArgThrArgLysTrg 119
               |||||
Db      866 CCGACGCGGAGTGCATACGTGCGCGCAGCGCCCGCTGACGAGGACGAGGACGCGCG 925
               |||||
QY      120 GlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
               |||||
Db      926 CAGGCGCTGGTGGATGCGCTTCCCGCGCTGGGCGGTGCTATTCATTATGTC----- 979
               |||||
QY      140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsn 159
               |||||
Db      980 ACGCGGAGTCCGATCCCTGTTCAGACCGACAGTGTGCGCGCCACGCAACCGACCCC 1039
               |||||
QY      160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGly 179
               |||||
Db      1040 GCCCGCCA-GCTGGCGCGGTGAAGAAGAGCGCGG-----GATCTGGC-GCA 1085
               |||||

```

```

QY 180 HsArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsnDheProGlnSerAsnLeu 199
    |||||
Db 1086 GGTGGCGAGCGC-----CTGGCGCCACCATCCAGCATATGAACAAGCGGTT 1133
QY 200 CysLeuLeuArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlu 219
    |||||
Db 1134 TGT-----CAGGCCGATPAGCCCTGCGCCGATCTTCGATACCC--CCGTGGCAG 1178
QY 220 LeuLysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeu 239
    |||||
Db 1179 GTGAGCAGACGAAAAAGTGGAGAGACCATAGCGGATGAGCCGTGATGCCAATATG 1238
QY 240 ThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlyTyrGly 257
    |||||
Db 1239 GTGGAGACGCTGCTGCTGCTGATGAACCTGCTCTCAGCCAGCTGGGCTGGGGC 1298
QY 258 ArgIleThrAspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyr 277
    |||||
Db 1299 AGATACCCAGGCGCAGGAGATACCGCCCTGCTGCTGCTTAAACGAAAACTACGAT 1358
QY 278 LeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
    |||||
Db 1359 CTGAGTACGATGTGTGTATACCGCCGCAAAACCGCGGTGCTGCTCAACGCTATG 1418
QY 298 MetAlaAlaLeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThr 317
    |||||
Db 1419 CTCGACGCGCTCAACCGCAGCATACGATGACGT-----ACGCTGGCTGCT 1463
QY 318 SerValLeuPheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlu 337
    |||||
Db 1464 GCT-----GCTGGCATGTGACACCAATATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
QY 338 LeuAsnTrpThrLeuProGly---GlnProAspAsnThrProProGlyGlyGluLeuVal 356
    |||||
Db 1512 TTATAGCTGACAGCTGCGCGGCTACACCGGGGAAATATCCCGCGCGGACAGCTGTGTG 1571
QY 357 PheGlnArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
    |||||
Db 1572 CTGGAGCGCTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1631
QY 377 GlnThrLeuGlnIleMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlu 396
    |||||
Db 1632 CAGGCGCTGACAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1691
QY 397 ValLysLeuThrLeuAlaGlnCysGlnGluArgAsnAlaGlnGlyMetCys-----413
    |||||
Db 1692 CAGGATGGCATAGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1751
QY 414 ---SerLeuAlaGlnPheThrGlnIleValAsnGlnAlaArgIleProAla 429
    |||||
Db 1752 GCGCGATATACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1802

```

RESULT 8

```

US-08-362-577C-23
: Sequence 23, Application US/08362577C
: Patent No. 5807673
: GENERAL INFORMATION:
: APPLICANT: Ohno, Tsuneya
: APPLICANT: Matsushima, Akio
: APPLICANT: Uehara, Hirotsugu
: APPLICANT: Eda, Soji
: TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
: NUMBER OF SEQUENCES: 25
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
: STREET: 6300 Sears Tower, 233 South Wacker Drive
: CITY: Chicago
: STATE: Illinois
: COUNTRY: United States of America
: ZIP: 60606-6402
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk

```

```

: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/362,577C
: FILING DATE: 27-MAR-1995
: CLASSIFICATION: 536
: ATTORNEY/AGENT INFORMATION:
: NAME: Rin-Laures, Li-Hsien
: REGISTRATION NUMBER: 33,547
: REFERENCE/DOCKET NUMBER: 19036/32420
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 312/474-6300
: TELEFAX: 312/474-0448
: INFORMATION FOR SEQ ID NO: 23:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5975 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: Genomic DNA
: ORIGINAL SOURCE:
: ORGANISM: Klebsiella pneumoniae
: STRAIN: Clinical isolate KI-50
: US-08-362-577C-23

Alignment Scores:
Pred. No.: 2,36e-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 1 Gaps: 14

US-09-777-566a-2 (1-440) x US-08-362-577C-23 (1-5975)
QY 2 LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrPro---GlnSerAla 20
    |||||
Db 578 CAGGGGCTGTATGACCTGTTATTCGCTGCGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 637
QY 21 PheAlaGlnSerGlnProGlnLeuLysLeuGlnSerValValIleValSerArgHisGly 40
    |||||
Db 638 GCGCGC-----GCGACCTGGCAGCTGAGAAATGTGATGACGACCCGACCGCGGT 668
QY 41 ValArgLapProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaTrp 59
    |||||
Db 689 ATTGCTCCGCGGACGCGCGGCGGCAACCGGAAAGCATGAGGCGCGGACCGGCGGCGGCGGCGG 748
QY 60 ProThrTrpProValLysLeuGlnTyrTrpLeuThrProArgGlnGlyGluLeuIleAlaTyr 79
    |||||
Db 749 ACCGAGGAGCACCCCATGAGCGGAGGAGCTCAGCGGCATGCTATGCGCGGCTGCTGCTGCTGCT 808
QY 80 LeuGlnHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysLysGlyCys 99
    |||||
Db 809 AAGGCGGCTGGAGGCGCAGCATATACCGCCAGCTGCGCTGCTG---CAGCGCGGATGC 865
QY 100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspLysLysLysLysLysLysLysLys 119
    |||||
Db 866 CCGACGCGGAGTCGATATACGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 925
QY 120 GlnAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
    |||||
Db 926 CAGGCGCTGTGATGAGCGCGCTTCCCGCGCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 979
QY 140 ThrSerSerProAspProLeuPheAsnProLeuLysThrGlyValLysGlnLeuAspAsn 159
    |||||
Db 980 AGCGGGATGCCATCCCTGTTTCAGACCGCAGCATGTCGCGCGCAGCAACCGGCGGCGGCGGCGG 1039
QY 160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGly 179
    |||||
Db 1040 GCCCGCCA-GCTGGCGGCTGAAAGAGAGAGGCGGG-----GATCTGGC-GCA 1085

```

180 HisArgGlnThrAlaPheArgGluLeuGlnArgValLeuAspPheProGlnSerAsnLeu 199
1086 GGTGGCGAGCGC-----CTGGCGCGGACCATTCACATATGTAACAGCGGTT 1133
200 CysLeuYsArgGluYsGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGlu 219
1134 TGT-----CAGCGCGATTAAGCCCTGCGATCTTCGATACC---CCGTGGCGAG 1178
220 LeuYsValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAspSerLeu 239
1179 GTCGAGCAGAGCAAAAGTGGCAAGCACCATTTACGAGCTAGAGCGTGCCCAATATG 1238
240 ThrGluIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlyTyrGly 257
1239 GTGAGAGCGTGGCTGGCTGGCTGAGTGAACAACTGCTTCAGCAGCAGCTGGCGGGCG 1298
258 ArgIleThrAspSerHisGlnThrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheYr 277
1299 AAGATCACCCAGGCGAGGAGATCACCCGCTGCGCGCTGTAAAGCAAACTACGAT 1358
278 LeuGlnArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
1359 CTGAGTACGATGTTTATATACCGCGCAAAACCGGGGTGGCTGCTCAACGCTATG 1418
298 MetAlaAlaLeuThrProHisProProGlnYsGlnAlaTyrGlyValThrLeuProThr 317
1419 CTCGACGCGCGCAAAACGAGGAGATGACGT-----ACGCTGGCTGCT 1463
318 SerValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyValAlaLeuGlu 337
1464 GCT-----GGTGGCATGACACCAATATGCGCATGCTGCGACGCGTGAATGAAAC 1511
338 LeuAsnThrThrLeuProGly---GlnProAspAsnThrProGlyGlyLeuVal 356
1512 TTTACTGTCGACCTCCCGGCTACAGCCGGGGAATATCCCGCGGAGCGAGCTGGG 1571
357 PheGluArgTyrPArgArgLeuSerAspAsnSerGlnThrIleGlnValSerLeuValPhe 376
1572 CTGGCGCGCTGGCGCAACGCGAAGAGCGAGAACCTATGCTGGGTGATTTCCAGGCG 1631
377 GlnThrLeuGlnGlnMetArgAspYsThrProLeuSerLeuAsnThrProProGlyGlu 396
1632 CAGGCGCTCGACGACCTGCTGCTGTCGACAGCGCGGACGCGACCCGATGCTGCGT 1691
397 ValYsLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCys----- 413
1692 CAGAGGTGCATCAGCCGGGCTGCGCTCAGACCCGATGCGGTGCTGCTCCAG 1751
414 ---SerLeuAlaGlyPheThrGlnIleValAsnGlnAlaArgIleProAla 429
1752 GCGGCTATACCGCCCTCGTCAGCGTATCGACGATCCGCGCCCGCGGCG 1802

RESULT 9
US-08-920-828-23
Sequence 23, Application US/08920828
Patent No. 5853398
GENERAL INFORMATION:
APPLICANT: Ohno, Tsuneya
APPLICANT: Matsuhisa, Akio
APPLICANT: Uehara, Hirotsugu
TITLE OF INVENTION: Probe for Diagnosing Infectious Disease
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun
STREET: 6300 Sears Tower, 233 South Wacker Drive
CITY: Chicago
STATE: Illinois
COUNTRY: United States of America
ZIP: 60606-6402
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/920,828
FILING DATE: 29-AUG-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/362,577
FILING DATE: 27-MAR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Rin-Laures, Li-Hsien
REGISTRATION NUMBER: 33,547
REFERENCE/DOCKET NUMBER: 19036/32420
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/474-6300
TELEFAX: 312/474-0448
TELEX: 25-3856
INFORMATION FOR SEQ ID NO: 23:
SEQUENCE CHARACTERISTICS:
LENGTH: 5975 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
ORIGINAL SOURCE:
ORGANISM: Klebsiella pneumoniae
STRAIN: Clinical Isolate KT-50
US-08-920-828-23

Alignment Scores:
Pred. No.: 2,366-27 Length: 5975
Score: 352.00 Matches: 116
Percent Similarity: 43.25% Conservative: 73
Best Local Similarity: 26.54% Mismatches: 211
Query Match: 15.29% Indels: 39
DB: 2 Gaps: 14

US-09-777-566a-2 (1-440) x US-08-920-828-23 (1-5975)

2 LysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrPro---GlnSerAla 20
578 CAGGGCGCTGTACGCGCTGTATCGCGCGGCTGCGCGCTGCGCGTGCATCTGCC 637
21 PheAlaGlnSerGluProGluLeuYsLeuGlnSerValAlaIleValSerArgHisGly 40
638 GCCGCC-----GCGACTGCGACGCTGGAGAAAGTGTGAGCTCAGCCGACAGGT 688
41 ValArgAlaProThrLysAlaThr---GlnLeuMetGlnAspValThrProAspAlaTrp 59
689 ATTCGTCGCGCAGCGCGCGCAACGCGGAGCGATCGAGCGCGCCAGCGCGGACCGTGG 748
60 ProThrTrpProValYsLeuGlyTyrPheLeuThrProArgGlyGlyGluLeuIleAlaYr 79
749 ACCGAGTGGACCAACCATGACGCGGAGCTCACCGGCGCATGCTATGCCGCTGGTCAAC 808
80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaYsGlyCys 99
809 AAAGGCGCTGGCGGAAGCCAGCATTAACCGCCAGCTCGGCTCTG---CAGCGCGATGC 865
100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspValAspValArgTyrGly 119
866 CCGAGCGCGGAGTGCATATACGTGGCGCGCACCGCGCTCAGCGGAGCGGAGCGAGCCGCC 925
120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
926 CAGGCGCTGCTGAGAGCGCTTCCCGGCTGCGCGCTGCTGCTATCATATATGTC----- 979
140 ThrSerSerProAspProLeuPheAsnProLeuLeuThrGlyValCysGlnLeuAspAsn 159
980 ACCGGGAGATCCGATCCCTGTTTCAGACCGCAAGTTCGCCGCCACGAAACCAACCC 1039
160 AlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGly 179

Db 1040 GCCCGCCA-GCTGGCCGGTGAAGAGAGCCGGG-----GATCGGC-GCA 1085
Qy 180 HsArgGlnThrAlaPheArgGlnLeuGluArgValLeuAsnProGlnSerAsnLeu 199
Db 1086 GGTGGGAGGCGG-----CTGGGCGGACATCCAGTATTAAGACGGCGGT 1133
Qy 200 CysLeuLysArgGlnLysGlnAspGlnSerCysSerLeuThrGlnAlaLeuProSerGln 219
Db 1134 TGT-----CAGCGCGATTAAGCCCTGCCCATCTTCATAC-----CCGTGGCAG 1178
Qy 220 LeuLysValSerAlaAspAsnValSerLeuThrGlnAlaValSerLeuAlaSerMetLeu 239
Db 1179 GTGACACAGACGAAAGTGGGAGACACCATTAAGCGAGCTGAGCGGATGCGCAATATG 1238
Qy 240 ThrGlnIlePheLeuLeuGlnGlnAlaGlnGlyMetPro-----GluProGlnTyrGly 257
Db 1239 GTGAGACAGCTGCGCTCGCGCTGGAGTGAACCTGCTCTGACGACAGCGGGGTGGGC 1298
Qy 258 ArgIleThrAspSerHisGlnIlePheAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyr 277
Db 1299 AAGATACCCAGGCGAGGACGATACCGCGCTGCTGCTTACGGAACATGAT 1358
Qy 278 LeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThrProLeuLeuAspLeuIle 297
Db 1359 CTGAGTAAGAGATGTTGTATACCGCGCAAAAGCGGGGTGCTGCTCAACGCTATG 1418
Qy 298 MetAlaAlaLeuThrProHisProGlnLysGlnAlaTyrGlyValThrLeuProThr 317
Db 1419 CTGACGCGGCTCAACCGGAGCATCAACGT-----ACGCTGGCGCT 1463
Qy 318 SerValLeuPheIleAlaGlnHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGln 337
Db 1464 GCT-----GGTGGCATGACACCAATATGCGTGGCGACAGCTGATGATAC 1511
Qy 338 LeuAsnTrpThrLeuProGln-----GlnProAspAsnThrProProGlnGlyGlnLeuVal 356
Db 1512 TTAGCTGGCAGTGGCGGCGCTACACCGGGGAAATATCCCGCGGACAGCACCCTGCTG 1571
Qy 357 PheGluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPhe 376
Db 1572 CTGAGCGCTGGCGGCAACGCGAGAGCGGAGAACGCTATCTGGGGCTATTTCCAGGCC 1631
Qy 377 GlnThrLeuGlnGlnMetArgAspLysTrpProLeuSerLeuAsnThrProProGlnGly 396
Db 1632 CAGGGCTCGACAGACCTGCTGCTGCAAGACCGCGGACGCGGATGCTGCTG 1691
Qy 397 ValLysLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCys----- 413
Db 1692 CAGAGTGGCATACCGCGGCTGCTGCTGACAGCGATGCTGCTGCTGCTGCTGCTGCTG 1751
Qy 414 ---SerLeuAlaGlyPheThrGlnIleValAsnGlnAlaArgIleProAla 429
Db 1752 GGGGCTATTACCGCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1802

RESULT 10
US-08-868-435-32
Sequence 32, Application US/08868435
Patent No. 6291221
GENERAL INFORMATION:
APPLICANT: Van Loon, Adolphus
APPLICANT: Mitchell, David
TITLE OF INVENTION: POLYPEPTIDES WITH PHRYASE ACTIVITY
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hoffmann-La Roche Inc.
STREET: 340 Kingsland Street
CITY: Nutley
STATE: New Jersey
COUNTRY: United States of America
ZIP: 07110
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/868,435
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/744,231
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Kass, Alan P
REGISTRATION NUMBER: 32142
REFERENCE/DOCKET NUMBER: Case Docket 9339
TELECOMMUNICATION INFORMATION:
TELEPHONE: (201) 235-4205
TELEFAX: (201) 235-2363
INFORMATION FOR SEQ. ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 1571 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: join(43..89, 147..1494)
FEATURE:
NAME/KEY: Intron
LOCATION: 90..146
FEATURE:
NAME/KEY: misc_feature
LOCATION: 894..999
OTHER INFORMATION: /note= "Position of PCR fragment"
US-08-868-435-32

Alignment Scores:
Pred. No.: 0.00419 Length: 1571
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.17% Indels: 114
Gaps: 24

US-09-777-566a-2 (1-440) x US-08-868-435-32 (1-1571)

Qy 7 ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro 26
Db 253 CCATCTTTTTCGCTCGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 303
Qy 27 GluLeuLysLeuLysSerValIleValSerArgHisGlyValArgAlaProThr 45
Db 304 GATTGCCGATCACTTGTGACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 363
Qy 46 ---LysAlaThrGlnLeuMetGlnAspValThrProAspAla----- 58
Db 364 TCCAGAGCAAAAGATAGAGAGCTTGTGACGCGCATCCAGCAATGCCACCGACTTC 423
Qy 59 -----TrpProThrTrpProValLysLeuGlyTrp-----LeuThr 70
Db 424 AAGGCAAGTTTCCTTTTGAAGAGCTATACCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 483
Qy 71 ProArgGlyGlyLeuLeuIleAlaTyrLeuGly-----HisTyrGlnArgGlnArgLeu 88
Db 484 CCGTTGGGAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
Qy 89 ValAlaAspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlnGlnValAlaIleIle 108
Db 544 CTGGCGCGAGTGTGTGGCG-----TTTATTTCCG 573
Qy 109 AlaAspValAspGlnArgThrArgLysTyrGlnGlnAlaPheAlaAlaGlyLeu----- 126

```

Db 574 GCGTCAGGCTCGACCGGGTTATGCTTCGGAGAGAACTTCATCGAGGGGTTCCAGCAG 633
QY 127 -----AlaProaspCysAlaIleThrVal 134
Db 634 GCGAGCTGGCTGATCCTGGCCGACGACCGCGCCGCTCCGGCATTTAGTGTGATTAT 693
QY 135 HisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuThrGlyVal 154
Db 694 ---CCGAGAGCGAGACGCTTCAAC-----AAVACGGTGTGACACACGGGTG 735
QY 155 CysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGlySerIle 174
Db 736 TGC-----ACGAAGTTTGAGCGAGTCGAGTGGAGATGAGGTT 774
QY 175 ---AlaAspPheThrGlyHisArgGlnThrAlaPheArgGlnLeuGlnArgValLeuAsn 193
Db 775 GCGGCCAATTTCAC-----GCGCTCTTT 798
QY 194 PheProGlnSerAsnLeuCysLeuLysArgGlnLysGlnAspGlnSerCysSerLeuThr 213
Db 799 GCACCC-----GACATCGAGCTCGCGCCGAGAGACATCTTCGCGGTGACGCGACAC 852
QY 214 GlnAlaLeuProSerGlnLeu-----LysValSerAlaAspAsnValSerLeuThrGly 231
Db 853 GACGAGGAGCTTGTCACTTAATGACATGCTGTTGATGACGTACGCGACGACAC 912
QY 232 AlaValSerLeuAlaSerMetLeuThrGlnIlePheLeuLeuGlnIleGlnIleMet 251
Db 913 GACGCAAGTCAGCTTCACCGCTTGTCACTTCACCTACACAAATGAG----- 960
QY 252 ProGlnProGlyTrpGlyArgIleThrAspSerHisGlnTrpAsnThrLeuSerLeu 271
Db 961 -----TGG-----AAGAAGTACAACTACTGACTGCTG 990
QY 272 HisAsnAlaGlnPheThrLeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThr 291
Db 991 -----GGCAAGTACTACGCTACGCGAGCGACCAACCCCTCGGACCGGCTCAGGGGATA 1044
QY 292 ProLeuLeuAspLeuIleMetAlaIleThrProHisProProGlnLysGln----- 309
Db 1045 GGGTTCACCAACGAGCTGATGCGCGGTGACTGCTCCCGCAGTCGAGGACACACAC 1104
QY 310 -----AlaTrpGlyValThrLeuProThrSerValLeuPheIleAla 323
Db 1105 ACTAAGTCGACTAGTCTCCACACCGCGCACCCTTCCTTGAACGCTACATGACGTC 1164
QY 324 -----GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTrpThr 341
Db 1165 GACTTTTTCACAGCAACAGCAATGGTTTCCATCTTCTTGCATTG----- 1209
QY 342 LeuProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgTrpArg 361
Db 1210 -----GGCCTGTACACGCGACTGACCCCTTGTCCCGACCTCGTGGAAGCGCCAG 1263
QY 362 ArgLeuSerAsp-----AsnSerGlnTrpIle-----GlnValSerLeuValPheGln 377
Db 1264 GAATTGGATGGGTAATTCATCTCGGCGGTGCTTCCTTGGCGCGGACCTACTTTCGAG 1323
QY 378 ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnVal 397
Db 1324 ACGATGCAATGCAATCGGAAAGAGCCTCTT---GTTGCGGCTTGTATTAATGACCGG 1380
QY 398 LysLeuThrLeuAlaGlyCysGlnGlnLysGlnAlaGlnIleMetCysSerLeuAlaGly 417
Db 1381 GTTGTGCACCTGCATCTGCGATGCTG---GACAAGCTGGGCGATGCAAGCTGAATGAC 1437
QY 418 PheThrGlnIleValAlaGlnLysVal 426
Db 1438 TTTGTCAAGGATTCAGTTGGGCGCAGA 1464

```

RESULT 11
 US-08-744-231-32
 ; Sequence 32, Application US/08744231

```

; Patent No. 6358722
; GENERAL INFORMATION:
; APPLICANT: Van Loon, Adolphus
; APPLICANT: Mitchell, David
; TITLE OF INVENTION: POLYPEPTIDES WITH PHYLASE ACTIVITY
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: United States of America
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/744,231
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,757
; FILING DATE: 18-APR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kass, Alan P
; REGISTRATION NUMBER: 32142
; REFERENCE/DOCKET NUMBER: Case Docket 9339
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-4205
; TELEFAX: (201) 235-2363
; INFORMATION FOR SEQ. ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1571 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: join(43..89, 147..1494)
; FEATURE:
; NAME/KEY: Intron
; LOCATION: 90..146
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 894..999
; OTHER INFORMATION: /note= "Position of PCR fragment"
; US-08-744-231-32

Alignment Scores:
Pred. No.: 0.00419 Length: 1571
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.17% Indels: 114
DB: 4 Gaps: 24

US-09-777-566a-2 (1-440) x US-08-744-231-32 (1-1571)
QY 7 ProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro 26
Db 253 CCATCTTTTCGCTCGAGCAGCAGCTGCTCGCTGCGAAGCTTCCCAAG----- 303
QY 27 GlnLeuLysLeuGlnLysValIleValSerArgHisGlyValArgAlaProThr----- 45
Db 304 GATTCGCGATCACCCTTGTGATCAGGTGCTATCGCGCCATGAGACCGGTAACCAACAC 363
QY 46 -----LysAlaThrGlnLeuMetGlnAspValThrProAspAla----- 58

```



```

Db      364  TCCAGAGCAAAAAGATATGAAGCTTGACGGCGATCCAGCCAAATGCCACCGACTTC 423
QY      59  -----TTPProthPrProValLysLeuGlyTrp-----LeuThr 70
Db      424  AAGGCAAGTTTCCTTTTGAAGACGTACATATCTGGGTGGGATGACCTACT 483
QY      71  ProArgGlyGlyLeuLeuLeuLeuLeuGly-----HisTyrGlnArgGlnArgLeu 88
Db      484  CCTTTGGGGAGAGAGCTGGTGAACCTCGCACTCAAGTTTCACCGAGGTACCAAGCT 543
QY      89  ValAlaAspLysLeuLeuAlaLysGlyCysProGlnSerGlyGlnAlaLeuLeu 108
Db      544  CTGGCCGCGAGTGTGGCGCG-----TTTATTCCG 573
QY      109  AlaAspValaAspLysArgThrArgLysThrGlyGlnAlaPheAlaAlaGlyLeu----- 126
Db      574  GCCTCAGAGCTCGACCGGGTTATGCTTCGGAGAGAAAGTTCAATCGAGGGGTTCACAG 633
QY      127  -----AlaProAspCysAlaLeuThrVal 134
Db      634  GCGAAGCTGGCTGATCTGGCGGAGACACCGCGCGCTCGCGATGATGATGATAT 693
QY      135  HisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThrGlyVal 154
Db      694  ---CCGGAGAGCGAGACTTCAAC-----AATACGCTGGACCAAGGTGTG 735
QY      155  CysGlnLeuAspAsnAlaAsnValThrAspAlaLeuSerArgAlaGlySerIle 174
Db      736  TGC-----ACGAAGTTTGAGCGGAGTCAAGCTGCGGAGATGAGAGTT 774
QY      175  ---AlaAspPheThrGlnHisArgGlnThrAlaPheArgLysLeuGlnArgValLeuAsn 193
Db      775  GCGGCCCAATTTCACT-----GGCCTCTTT 798
QY      194  PheProGlnSerAsnLeuLysArgGlyLysGlnAspLysSerSerLeuThr 213
Db      799  GCACCC-----GACATCCGAGCTCGCGCGGAGAACATCTTCTCGCGTGAAGCTTACA 852
QY      214  GlnAlaLeuProSerGlyLeu-----LysValSerAlaAspAsnValSerLeuThrGly 231
Db      853  GACGAGAGCTTGTCACTGAATGACATGTCTTGTATGACGAGCGGACGACGAGC 912
QY      232  AlaValSerLeuLysSerMetLeuThrGlnIlePheLeuLeuGlnAlaGlnGlyMet 251
Db      913  GACGCAAGTCACTGTACCCGTTCTGTCACTTCACTCAATGAG----- 960
QY      252  ProGlnProGlyTyrPheArgLysLeuThrAspSerHisGlnThrAsnThrLeuSerLeu 271
Db      961  -----TGG-----AAGAAGTACAACTACCTTCACTGCTTG 990
QY      272  HisAsnAlaGlnPheTyrLeuLeuGlnArgThrProGlnValAlaArgSerArgAlaThr 291
Db      991  -----GCGAAGTACTACGCTGACGGCGGAGCAACCCCTGGAGCGGCTGAGGGATA 1044
QY      292  ProLeuLeuAspLeuLeuMetAlaAlaLeuThrProHisProProGlnLysGln----- 309
Db      1045  GGGTTCACCAACAGAGTGTATGCCGGTGTGATCGTTCGCCAGTGAAGGACACACAGC 1104
QY      310  -----AlaTyrGlyValThrLeuProThrSerValLeuPheIleAla 323
Db      1105  ACTAAGCTGACTTACTTCCAAACCGGCACTTCCCGTGAACGCTACCATGTACGTC 1164
QY      324  -----GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeuAsnTyrThr 341
Db      1165  GACTTTTCACACGACAAAGCATGTTTCCATCTCTTTCGACTG----- 1209
QY      342  LeuProGlyGlnProAspAsnThrProProGlyGlyGlyLeuValPheGlnArgTyrArg 361
Db      1210  -----GGCCTGTACAAAGCAAGCTGAACCTTGTCCGAGACTCGGTGGAAACGCGCAAG 1263
QY      362  ArgLeuSerAsp-----AsnSerGlnTyrPhe-----GlnValSerLeuValPheGln 377
Db      1264  GAATTGGATGGGATATTCTGCATCTGGGTGGTGGCTTTGGCGCGGAGCGCTACTTTCGAG 1323

```

```

QY      378  ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlnVal 397
Db      1324  ACGATGCAATGCAAGCTGCGAAAGAGCACTCTT---CTTGCGCGCTTGATTAATGACCG 1380
QY      398  LysLeuThrLeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGly 417
Db      1381  GTTGCGCACTGATGAGGTGCGATGCTGATGTG---GACAAGCTGGGCGGATGCAAGCTGAATGAC 1437
QY      418  PheThrGlnIleValAsnGlnAlaArg 426
Db      1438  TTGTCTCAAGGATTTGAGTTGGCCACA 1464

```

```

RESULT 12
US-09-044-718-10
: Sequence 10, Application US/09044718
: Patent No. 6391605
: GENERAL INFORMATION:
: APPLICANT: KOSTREMA, Dirk
: APPLICANT: PASAMONTES, Luis
: APPLICANT: TOMSCHY, Andrea
: APPLICANT: van LOON, Adolphus
: APPLICANT: VOGEL, Kurt
: APPLICANT: WYSS, Markus
: TITLE OF INVENTION: MODIFIED PHYTASES
: FILE REFERENCE: Modified Phytases
: CURRENT APPLICATION NUMBER: US/09/044,718
: CURRENT FILING DATE: 1998-03-19
: PRIOR APPLICATION NUMBER: EP 97810175.6
: PRIOR FILING DATE: 1997-03-25
: NUMBER OF SEQ ID NOS: 82
: SOFTWARE: Patentln Ver. 2.1
: SEQ ID NO 10
: LENGTH: 1571
: TYPE: DNA
: ORGANISM: Aspergillus fumigatus
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (43)..(90)
: NAME/KEY: CDS
: LOCATION: (148)..(1494)
US-09-044-718-10

Alignment Scores:
Pred. No.: 0.00419 Length: 1571
Score: 119.00 Matches: 102
Percent Similarity: 36.67% Conservative: 70
Best Local Similarity: 21.75% Mismatches: 183
Query Match: 5.17% Indels: 114
DB: 4 Gaps: 24

US-09-777-566a-2 (1-440) x US-09-044-718-10 (1-1571)
QY      7  ProPheLeuSerLeuLeuLeuProLeuThrProGlnSerAlaPheAlaGlnSerGluPro 26
Db      253  CCATTCCTTTTCGCTCGAGAGCAAGCTGTCCGTCTCGAGTAAAGCTTCCCAAG----- 303
QY      27  GlnLeuLysLeuGlnSerValAlaIleValSerArgHisGlyValArgAlaProThr--- 45
Db      304  GATTGCCGATCACTTGTGTACAGGTCTATCCGCCATGTGAGCGCGGTACCCACAGC 363
QY      46  -----LysAlaThrGlnLeuMetGlnAspValThrProAspAla----- 58
Db      364  TCCAGAGCAAAAAGATATGAAGCTTGACGGCGATCCAGCCAAATGCCACCGACTTC 423
QY      59  -----TTPProthPrProValLysLeuGlyTrp-----LeuThr 70
Db      424  AAGGCAAGTTTCCTTTTGAAGACGTACATATCTGGGTGGGATGACCTACT 483
QY      71  ProArgGlyGlyLeuLeuLeuLeuLeuGly-----HisTyrGlnArgGlnArgLeu 88
Db      484  CCTTTGGGGAGAGAGCTGGTGAACCTCGCACTCAAGTTTCACCGAGGTACCAAGCT 543

```



```
Db 156 AGTCCATT-----GACACCTTTCCACCTGAC 182
Oy 62 -----TriproValLysLeuGlyTyrLeuThrProArgLysGlyLeu 75
Db 183 CCCATTAAGAAATCCTCATGCGCACAGAGATTGGCCAACTACACCAGCTGGCATGGAG 242
Oy 76 LeuIleAlaTyrLeuGlyHisTyrGlnArgGlnArgLeuValAlaAspLysLeuAla 95
Db 243 CAGCATTTAGAACTTGGAGAGTATATAGAAAGAGATAT----- 281
Oy 96 LysLysGlyCysProGlnSer-----GlyGlnValAlaIleIleAlaLysValAsp 112
Db 282 AAAAAATTCCTTGAAATGAGCTCTATTAACAATGAACAGATTATATTCGAGACAGACGTT 341
Oy 113 GluArgThrArgLysThrGlyLysAlaPheAlaAlaGlyLeuAlaProAspCysAlaIle 132
Db 342 GACCGGACTTTCATGATGCTCTATGACAAACCTGGACGCCCTTTCCGCCAAGAGTGC 401
Oy 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db 402 AGCATCTGG-----AATCCTATTCCTACTGCGACGCCCATCCCGGTG 443
Oy 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGly 172
Db 444 CACACAGTTCCTCTTCTGAA----- 464
Oy 173 SerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgLysLeuGlnArgValLeu 192
Db 465 -----CATCAGTTGCTA 476
Oy 193 AsnProGlnSerAsnLeuLysLysArgLysGlnAspLysSerLysSerLeu 212
Db 477 TACCTGACCTTCAGGAAAC---TGCCCTCGCTTTCAAGAACTTGAGAGTGAAGCTTTGAAA 533
Oy 213 ThrGlnAlaLeuProSerGlnLeuLysValSerAlaAspValSerLeuThrGlyAla 232
Db 534 TCAGAGGAATTCACAGAGAGCTGCACCTTATAGGATTTATAGCTACCTTTGGGAAA 593
Oy 233 ValSer-----LeuAlaSerMetLeuThrGlnIlePhe 243
Db 594 CTTTCAGATTCATAGGCCAGGACCTTTTGGAAATTTGGAGTAAAGTTCACGACCCCTTA 653
Oy 244 LeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThrAspSerHis 263
Db 654 TATTGTGACAGTTCACAAATTCCTACTTACCTCCGCGCCACGACGACACCATGACT 713
Oy 264 GlnTyrPasnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGlnArgThrPro 283
Db 714 AGTTGACAGAAATTTGCAGAAATTTGCCCTCTGCTCTATGAGAAATTCACAAAGCAAAA 773
Oy 284 GluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrPro 303
Db 774 GAGAAATCTAGGCTCCAAAGGGGTGTCCTGTCATGAATCCTCAATCAGATG----- 827
Oy 304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAla 323
Db 828 -----AAGGAGACCAACTCAGATACCAAGCTACAAAACCTTATCATGTATCT 875
Oy 324 GlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlnLeu---AsnThrPheLeu 342
Db 876 GCGCATGACACTACTGTGAGTGGCTACAGATGGCGTGAAGATTTACCAAGCACTCTT 935
Oy 343 ProGlyGlnProAspAsnThrProProGlyGlyGlnLeuValPheGlnArgThrPargArg 362
Db 936 CCT-----CCCTATGCTTCTTGCCACTGACGGAATTTACTTTAGAAAGGGGAGTAC 989
Oy 363 LeuSerAspAsnSerGlnTyrIleGlnValSerLeuValPheGlnThrLeuGlnGlnMet 382
Db 990 TTT-----GTGAGATGTACTATCGAAT----- 1013
Oy 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThrLeuAla 402
Db 1014 -----GAGACGACAGACGACCGCTATCCCTCATGTACCT 1049
```

```
Oy 403 GLYCYSGLUGLUARGASNAAGLNGLYMETCYSSERLEUALAGLYPHERTHRGLNILEVAL 422
Db 1050 GGCTGACGCCCT-----ACGTCTCCTCTGAGAGGTTTCTGAGCTGGTT 1094

RESULT 14
US-08-579-823A-1
; Sequence 1, Application US/08579823A
; Patent No. 6080409
; GENERAL INFORMATION:
; APPLICANT: Laus, Reiner
; APPLICANT: Ruegg, Curtis L.
; TITLE OF INVENTION: Immunostimulatory Composition and Method
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Delinger & Associates
; STREET: 350 Cambridge Ave. Suite 250
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94306
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/579,823A
; FILING DATE: 03-DEC-1998
; CLASSIFICATION: 336
; ATTORNEY/AGENT INFORMATION:
; NAME: Judge, Linda R.
; REGISTRATION NUMBER: 42,702
; REFERENCE/DOCKET NUMBER: 7636-0010
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-324-0880
; TELEFAX: 650-324-0960
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1588 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; INDIVIDUAL ISOLATE: Prostatic acid phosphatase-GM-CSF
; IMMEDIATE SOURCE:
; LIBRARY: prostate carcinoma Lncap.FGC; PBMC

US-08-579-823A-1

Alignment Scores:
Pred. No.: 0.00702 Length: 1588
Score: 117.00 Matches: 87
Percent Similarity: 34.09% Conservative: 63
Best Local Similarity: 19.77% Mismatches: 178
Query Match: 5.08% Indels: 112
DB: 3 Gaps: 15

US-09-777-566A-2 (1-440) x US-08-579-823A-1 (1-1588)

Oy 3 AlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAla 22
Db 51 AGCCTTAGCCTTGCGCTTCTTCTTTCTTTCTTGCGCTAGACCGAGGTGACTAGCC 110
Oy 23 GlnSerGlnProGlnLeuLysLeuGlnSerValAlaIleValSerArgHisGlyValArg 42
Db 111 AAG-----GAGTTGAGATTGTGTGACTTGTGTGTTTGGCATGGACGACCA 155
```


Db 51 AGCCTTAGCCTGGCTTCCTGTTCTTTTCTGCTTTTCTGCTAGACCGAAGTACTAGCC 110
QY 23 GlnSerGluProGlnLeuLysLeuGlnSerValIleValSerArgHisGlyValArg 42
Db 111 AAG-----GAGTTGAGATTGTGACTTGTGGTGGTTCGGATGGAGACCGA 155
QY 43 AlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThr 61
Db 156 AGTCCCATTT-----GACACCTTTTCCCACTGAC 182
QY 62 -----TrpProValLysLeuGlyTyrPheThrProArgGlyGly 75
Db 183 CCCATAAAGGAATCCCATGAGCCACAGAGATTGGCCAACTCCACCAAGCTGGGCAATGAG 242
QY 76 LeuIleAlaTyrLeuGlnLysIleTyrGlnArgGlnArgLeuValAlaAspGlyLeuAla 95
Db 243 CAGCATTATGAACTTGGAGAGATATATAGAAAGAGATAT----- 281
QY 96 LysLysGlyCysProGlnSer-----GlyGlnValAlaIleIleAlaAspValAsp 112
Db 282 AGAAATTTCTTGAAATGAGTCCATATAAACAATGAACAGCTTATATTCGAAGCACAGAGCTT 341
QY 113 GluArgThrArgLysThrGlyGluAlaPheAlaIleGlyLeuAlaProAspCysAlaIle 132
Db 342 GACCGGACTTTGATGAGTCTATGACAAACCTGGCAGCCCTGTTTCCCCAGAGAGTGC 401
QY 133 ThrValHisThrGlnAlaAspThrSerSerProAspProLeuPheAsnProLeuLysThr 152
Db 402 AGCATCTGG-----AATCCTATCTCTGAGCAGCCCATCCGGTG 443
QY 153 GlyValCysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAlaGly 172
Db 444 CACACAGTTCTCTTTCTGAA----- 464
QY 173 SerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeu 192
Db 465 -----GATCAGTTGCTA 476
QY 193 AsnPheProGlnSerAsnLeuLysLeuLysArgGluLysGlnAspGluSerCysSerLeu 212
Db 477 TACCTGCCCTTCAGGAACT---TGCCCTCGTTTCAGAACTTCAGAGAGTGAAGCTTTGAAA 533
QY 213 ThrGlnAlaLeuProSerGluLeuLysValSerAlaAspAsnValSerLeuThrGlyAla 232
Db 534 TCAGAGGAATTCAGAGAGCTGCACCTTAAAGATTATAGCTACTCTTGCGGAAA 593
QY 233 ValSer-----LeuAlaSerMetLeuThrGluIlePhe 243
Db 594 CTTTCAGGATTCATGCGCCAGACCTTTTGGAAATTGGAGTAAAGTCAACGACCTTA 653
QY 244 LeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThrAspSerHis 263
Db 654 TATTGGAGAGTGTTCACAAATTTCACCTTACCTCCCTGGCCACCTGAGACACCATGACT 713
QY 264 GlnTyrAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgThrPro 283
Db 714 AAGTTGAGAGAAATTGTCAGAAATTGTCCCTGTCCTATGAGAAATTCACAAAGAGAAA 773
QY 284 GluValAlaArgSerArgIleThrProLeuLeuAspLeuIleMetAlaIleAlaLeuThrPro 303
Db 774 GAGAAATCTAGGCTCCCAAGGGGTGCTCCGTCATGAATCTCTCAATCAGATG----- 827
QY 304 HisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAla 323
Db 828 -----AAGAGAGCACTCAGATACCAAGCTACAAAACCTATATCATGTATCT 875
QY 324 GlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGlnLeu---AsnTyrThrLeu 342
Db 876 GCGCATGACACTACTGAGTGCTGCTACAGATGGCGCTAGATGTTTACAAAGACATCCTT 935
QY 343 ProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgThrArgArg 362
Db 936 CCT-----CCTATGCTTCTTGCCACTTGACGAGAAATGTACTTGAAGAGGGGAGTAC 989

QY 363 LeuSerAspAsnSerGlnThrPileGlnValSerLeuValPheGlnThrLeuGlnMet 382
Db 990 TTT-----GTGAGATGATACTATCGAAT----- 1013
QY 383 ArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThrLeuAla 402
Db 1014 -----GAGACGACAGCAGCCGATACCCCTCATGCTTA 1049
QY 403 GlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIleVal 422
Db 1050 GGCTGCAAGCCT-----AGCTGTCTCTGAGAGAGTTTGTCTGAGCTGTT 1094

Search completed: June 12, 2003, 12:11:09
Job time : 124 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: June 12, 2003, 11:33:34 ; Search time 226 Seconds
(without alignments)
2724.763 Million cell updates/sec

Title: US-09-777-566a-2

Perfect score: 2302

Sequence: 1 MKAILIFLSILPLTPQSA.....IVNEARIPACSLRSHHHHH 440

Scoring table:

	BLOSUM62
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 870385 segs, 699768693 residues

Total number of hits satisfying chosen parameters: 1740770

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL-frame+ p2n model -DEV-xlh
-Q=/cgn2.1/USPFO/spool/US09777566/Runat_05062003.154806.6495/app.query.fasta_1.583
-DB-published.Applifications_NA -OFMT-fastlap -SUFFIX-rnpb -MINMATCH-0.1
-LOOPEXT-0 -LOOPEXT-0 -UNITS-bits -START-1 -END-1 -MATRIX-Blosum62
-TRANS-human40.cdi -LIST-45 -DOCALLIGN-200 -THR_SCORE-pct -THR_MAX-100
-THR_MIN-0 -ALIGN-15 -MODE-LOCAL -OUTFMT-pto -NORM-ext -HEAPSIZE-500 -MINLEN-0
-MAXLEN-2000000000 -USER-US09777566.ecgn1.1.146 @runat_05062003.154806.6495
-NCPU-6 -ICPU-3 -NO_MMAP -LARGEQUERY -NEG_SCORES-0 -WAIT -DSPBLOCK-100
-LONGLOG -DEV_TIMEOUT-120 -WARN_TIMEOUT-30 -THREADS-1 -XGAPOP-10 -XGAPEXT-0.5
-FGAPOP-6 -FGAPEXT-7 -YGAPOP-10 -YGAPEXT-0.5 -DELOP-6 -DELEXT-7

Database : Published Applifications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC1_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCUT5_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	2302	100.0	1323	US-10-034-985-1
2	2302	100.0	1323	US-09-777-566a-1
3	2302	100.0	1323	US-09-866-379-1
4	2235	97.1	1901	US-09-866-379-7

	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
	2226	96.7	1901	10	US-09-866-379-9	Sequence 9, Appl1																																			
	2221	96.5	1489	9	US-10-266-041-9	Sequence 6, Appl1																																			
	2221	96.5	1901	10	US-09-866-379-6	Sequence 9, Appl1																																			
	2220	96.4	1901	10	US-09-866-379-5	Sequence 5, Appl1																																			
	960.5	41.7	1326	9	US-10-021-723A-11	Sequence 11, Appl1																																			
	938.5	40.8	1325	9	US-10-021-723A-3	Sequence 3, Appl1																																			
	918.5	39.9	1325	9	US-10-021-723A-9	Sequence 9, Appl1																																			
	760	33.0	1431	9	US-10-021-723A-1	Sequence 1, Appl1																																			
	476.5	20.7	1230	9	US-10-021-723A-5	Sequence 7, Appl1																																			
	472.5	20.5	1266	9	US-10-021-723A-5	Sequence 7, Appl1																																			
	119	5.2	1398	9	US-09-803-454-19	Sequence 19, Appl1																																			
	119	5.2	1425	9	US-10-213-990-23	Sequence 23, Appl1																																			
	119	5.2	1455	9	US-10-213-990-22	Sequence 22, Appl1																																			
	119	5.2	1571	9	US-10-062-848-10	Sequence 10, Appl1																																			
	117	5.1	3061	9	US-09-974-546-47	Sequence 47, Appl1																																			
	115.5	5.0	1372	10	US-09-765-111A-13	Sequence 13, Appl1																																			
	115.5	5.0	2707	9	US-10-097-340-242	Sequence 242, App																																			
	114.5	5.0	1758	9	US-09-895-793-979	Sequence 979, App																																			
	114.5	5.0	1758	10	US-09-822-827-979	Sequence 979, App																																			
	113.5	4.9	2625	10	US-09-765-111A-5	Sequence 5, Appl1																																			
	113.5	4.9	2711	10	US-09-765-111A-24	Sequence 24, Appl1																																			
	110.5	4.8	2406	9	US-09-895-793-980	Sequence 980, App																																			
	110.5	4.8	2406	10	US-09-822-827-980	Sequence 980, App																																			
	107	4.6	1431	9	US-09-738-626-1548	Sequence 1548, Ap																																			
	107	4.6	3309400	9	US-09-738-626-1	Sequence 216, App																																			
	106.5	4.6	3147	10	US-09-741-669-216	Sequence 216, App																																			
	106.5	4.6	3147	10	US-09-815-242-5965	Sequence 5965, Ap																																			
	106	4.6	2105	9	US-10-153-668-30	Sequence 30, Appl1																																			
	105	4.6	64	9	US-10-021-723A-23	Sequence 23, Appl1																																			
	105	4.6	3097	10	US-09-962-832-227	Sequence 227, App																																			
	103.5	4.5	3309400	9	US-09-738-626-1	Sequence 1, Appl1																																			
	103	4.5	536165	9	US-09-939-964-1	Sequence 1, Appl1																																			
	102	4.4	1725	9	US-09-738-626-763	Sequence 763, App																																			
	101	4.4	3804	9	US-09-915-814-3	Sequence 3, Appl1																																			
	101	4.4	81940	9	US-09-759-508B-1	Sequence 1, Appl1																																			
	100	4.3	1803	9	US-10-072-307-57	Sequence 57, Appl1																																			
	100	4.3	1971	10	US-09-974-300-322	Sequence 322, App																																			
	99.5	4.3	1845	9	US-10-062-848-7	Sequence 7, Appl1																																			
	99	4.3	1803	9	US-10-072-307-19	Sequence 19, Appl1																																			
	99	4.3	1615	9	US-10-108-605-302	Sequence 302, App																																			
	98.5	4.3	14756	9	US-10-160-758-5	Sequence 5, Appl1																																			

ALIGNMENTS

RESULT 1
US-10-034-985-1
; Sequence 1, Application US/10034985
; Publication No. US20030049815A1
; GENERAL INFORMATION:
; APPLICANT: Kreitz, Keith
; TITLE OF INVENTION: NOVEL PHRASE
; FILE REFERENCE: 09010/029003
; CURRENT APPLICATION NUMBER: US/10/034,985
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: US/09/580,515
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 08/910,798
; PRIOR FILING DATE: 1997-08-13
; PRIOR APPLICATION NUMBER: 09/259,214
; PRIOR FILING DATE: 1999-03-01
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1320)
; NAME/KEY: misc_feature

LOCATION: (1)...(1323)
OTHER INFORMATION: n = A,T,C or G
US-10-034-985-1

Alignment Scores:

Pred. No.:	4.1e-264	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	9	Gaps:	0

US-09-777-566a-2 (1-440) x US-10-034-985-1 (1-1323)

```

QY      1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
DB      1 ATGAAAGCATCTTAATCCATTATTTATCTCTTCTATTCGGTTAACCCCGCATGTGCA 60
QY      21 PheAlaGlnSerGluProGluLeuLysLeuGluSerValValIleValSerArgHisGly 40
DB      61 TTCGCTCAGAGTAGAGCCGAGCTGAAGTGAAGTGGTGTGATGTCTCAGTGGTATGGT 120
QY      41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValAlaThrProAspAlaTrpPro 60
DB      121 GTGCGTGTCTCAACCAAGGCGACGACGATGATGACAGATGTCACCCGACAGCATGGCCA 180
QY      61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTrpLeu 80
DB      181 ACCTGGCCGGTAAACTGGGTGGTGGTGCACACCGCGNGGTGAGTAAATCCGCTATCTC 240
QY      81 GlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
DB      241 GGACATTAACCAAGCGCGCTGTGTAAGCGAGATTCCTGGCGAAAAAGGCTTCGCCG 300
QY      101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGlnArgThrArgLysThrGlnAla 120
DB      301 CAGTGTGTGTCAGTCCGATTTATGTCTGATGTCGACGACGATACCCGTTAAACAGCGCA 360
QY      121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
DB      361 GCCTTCGCGCGCGGGGTGCACCTGACTGTCATTAACCGTATACATCCAGGAGATAG 420
QY      141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
DB      421 TTCAGTCCCGATCCGTTATTAATCCTCAAAAACGGGGTGGCCAACTGGTAAACGGG 480
QY      161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
DB      481 AACGTGACTGACGCGCATCTCTACAGCAGGCGAGGAGGTCAATGCTTACCGGGCAT 540
QY      181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
DB      541 CGGCAAAACGGCGTTTCGCAACTGCAAGCGGTGCTTAATTTCCGCAATCAAACTGTGC 600
QY      201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
DB      601 CTTAAACGTAAGCAAGACGAGCAAGCTGTTCAATTAACGACGATTAACATTCGGAATC 660
QY      221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
DB      661 AAGGTGAGCGCGGACAAATGTCATTAACCGGTGCGGTAAAGCTCCATCAATGCTGACG 720
QY      241 GlnIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpPheArgIleThr 260
DB      721 GAGATATTTCTCTCTCAACACACAGGAAATGCCGAGCGGGGTGGGAGGATGATCACC 780
QY      261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
DB      781 GATTTCACACCACTGGAACACCTTGTAACTTGCATTAACGCCCAATTAATTTGCTACAA 840
QY      281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
DB      841 CGCAGCGCCAGAGGTGGCCGCGACGCGCGCCGCTTAATTTGATTCATGGCAGCG 900

```

```

QY      301 LeuThrProHisProProGlnLysGlnAlaTrpGlyValIleThrLeuProThrSerValLeu 320
DB      901 TTACACCCCATCCATCCACCGCAAAAACAGCGCTATGAGTGTACATTTACCATTTAGTACTG 960
QY      321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
DB      961 TTTATTCGCGGACACGATTAATCTGGCAAAATCTGGGGCGGCGCACTGAGCTCAACTGG 1020
QY      341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
DB      1021 ACCGTTCCGCTAGCGGATTAACACGCGCGCGGTGTCACTGGTGTGAACCTGG 1080
QY      361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
DB      1081 CGTCGGCTAAGCATTAACACGCGCATGATTCAGGTTTCGTGCTTCCAGCATTTACAG 1140
QY      381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
DB      1141 CAGATCGGTGATTAACGCGCGCTGCTATTAAATACCGCCCGGAGAGGTGAACCTGACC 1200
QY      401 LeuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
DB      1201 CTGGCAGGATGTAAAGCGAAATGCGCAGGGCATGTGTGTTGCGCAGGTTTACGCAA 1260
QY      421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHisHis 440
DB      1261 ATCGTGAATGAACACGACGATACCGGGGTGCGATTTGATTCATCATCCATCACTAC 1320

```

RESULT 2

```

US-09-777-566a-1
; Sequence 1, Application US/09777566a
; Patent No. US20010055788a1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVER1370-6
; CURRENT APPLICATION NUMBER: US/09/777,566A
; CURRENT FILING DATE: 2001-06-11
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1320)
; NAME/KEY: misc feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
US-09-777-566a-1

```

Alignment Scores:

Pred. No.:	4.1e-264	Length:	1323
Score:	2302.00	Matches:	440
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	10	Gaps:	0

US-09-777-566A-2 (1-440) x US-09-777-566A-1 (1-1323)

```

QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGATTCGGTTAAACCCCGAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCCTCAGAGTGGCCGAGCTGAAAGTGGAAAGTGGTGAATGTCAGTCGTCATGCT 120
QY 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 121 GTGGGTGCTCCACACCAAGCCACGCACTGATGCGAGATGTCACCCCGACGCAATGGCCA 180
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACTGGTGGTGGTACACCGCGGNGTGGAGACTGAATGCGCTATGCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTAACCAACCCACGCTGCTGGTACCGACGATTTGCGGCAAAAAGGCGTGGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAAGTCCGATTAATTCCTATGTCACGAGCGTACCCGTTAAACAGCGCA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCGTTCGGCCGCGGGCTGGACCTGACCTGCAATTAACGTAACACCAAGCAATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAAATCCTTAAAAACTGCGTTGGCACTGATTAACG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACCGGATCTCAGCAGGCGAGAGGGTCAATTCGCTTAACCGGGCT 540
QY 181 ArgGlnThrAlaPheArgGluLeuGlnuArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db 541 CGGCAAAAGCGGTTTCGGGAACCTGGAACGGGCTTAATTTCCGCATCAACCTTGTC 600
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACCTGAGAAACAGACGAAAGCTGTTCAATTAACGACGCAATTAACGACG 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGTGAGCGCGCAATGCTCATTAAACGGTGGTGAAGCTGCAATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAAGCAACAGGGAATCCGCGAGCCGGGTGGGGAAGATCACC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACACGAGTGAACACCTGCTAGTTGCAATACGCGCAATTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCCAGAGGTTGGCCCGAGCGCGCACCCCGTTATGGAATTTGATCTGACAGC 900
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATCCACCGCAAAACAGCGCGTATGTTGACATTACCCACTTCATGACTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCGGACAGACTAATAATCTGGCAAAATCTCGGCGGCGCACTGAGACTCAACTG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheLysArgTrp 360
Db 1021 ACGCTTCCCGGTACACCGGATTAACGCGCGGAGTGTGAACCTGCTTTGTAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

```

```

Db 1081 CGTCGCTAAGGATTAACAGCCAGTGGATTCAGTTGCTGCTTCACAGACTTACAG 1140
QY 381 GlnMetAlaAspLysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThr 400
Db 1141 CAGATGCGGTGAATAAACCGCGCTGTCATTAAATACCGCCCGGAGAGTGAACGTGACC 1200
QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1201 CTGGCAGAGTGTGAAGAGCGAAATGCGCAGGCGATGTGCTGTGGCAGTTTACGCA 1260
QY 421 IleValAsnGluAlaArgIleProAlaCysSerLeuArgSerHisHisHisHis 440
Db 1261 ATCGTGAATGAAGCAGCATACCGGCGTGTGATTTGAGATCTCATCACCATCATCAC 1320

RESULT 3
US-09-866-379-1
; Sequence 1, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
; FILE REFERENCE: DIVERI370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIOR FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 1323
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(1323)
; OTHER INFORMATION: n is any nucleotide
; NAME/KEY: CDS
; LOCATION: (1)..(1323)
; OTHER INFORMATION:
US-09-866-379-1

Alignment Scores:
Pred. No.: 4,1e-264 Length: 1323
Score: 2302.00 Matches: 440
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-866-379-1 (1-1323)
QY 1 MetLysAlaIleLeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 1 ATGAAGGAGATCTTAATCCATTTTATCTCTTCGATTCGGTTAAACCCCGAATCTGCA 60
QY 21 PheAlaGlnSerGluProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
Db 61 TTCCTCAGAGTGGCCGAGCTGAAAGTGGAAAGTGGTGAATGTCAGTCGTCATGCT 120
QY 41 ValArgAlaProThrLysAlaIleThrGlnLeuMetGlnAspValThrProAspAlaThrPro 60
Db 121 GTGGGTGCTCCACACCAAGCCACGCACTGATGCGAGATGTCACCCCGACGCAATGGCCA 1260
QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyGluLeuIleAlaTyrLeu 80
Db 181 ACCTGGCCGGTAAACTGGTGGTGGTACACCGCGGNGTGGAGACTGAATGCGCTATGCTC 240
QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 241 GGACATTAACCAACCCACGCTGCTGGTACCGACGATTTGCGGCAAAAAGGCGTGGCCG 300
QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGlyGlu 120
Db 301 CAGTCTGGTCAAGTCCGATTAATTCCTATGTCACGAGCGTACCCGTTAAACAGCGCA 360
QY 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 361 GCGTTCGGCCGCGGGCTGGACCTGACCTGCAATTAACGTAACACCAAGCAATACG 420
QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
Db 421 TCCAGTCCCGATCCGTTATTTAAATCCTTAAAAACTGCGTTGGCACTGATTAACG 480
QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 481 AACGTGACTGACCGGATCTCAGCAGGCGAGAGGGTCAATTCGCTTAACCGGGCT 540
QY 181 ArgGlnThrAlaPheArgGluLeuGlnuArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db 541 CGGCAAAAGCGGTTTCGGGAACCTGGAACGGGCTTAATTTCCGCATCAACCTTGTC 600
QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 601 CTTAAACCTGAGAAACAGACGAAAGCTGTTCAATTAACGACGCAATTAACGACG 660
QY 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 661 AAGTGAGCGCGCAATGCTCATTAAACGGTGGTGAAGCTGCAATCAATGCTGACG 720
QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 721 GAGATATTTCTCTGCAACAAGCAACAGGGAATCCGCGAGCCGGGTGGGGAAGATCACC 780
QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGln 280
Db 781 GATTCACACACGAGTGAACACCTGCTAGTTGCAATACGCGCAATTTATTTGCTACAA 840
QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db 841 CGCAGCGCCAGAGGTTGGCCCGAGCGCGCACCCCGTTATGGAATTTGATCTGACAGC 900
QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 901 TTGACGCCCATCCACCGCAAAACAGCGCGTATGTTGACATTACCCACTTCATGACTG 960
QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnTrp 340
Db 961 TTTATTTGCGGACAGACTAATAATCTGGCAAAATCTCGGCGGCGCACTGAGACTCAACTG 1020
QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheLysArgTrp 360
Db 1021 ACGCTTCCCGGTACACCGGATTAACGCGCGGAGTGTGAACCTGCTTTGTAACGCTGG 1080
QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380

```

QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 |||||
 Db 121 GTGCGTGTCTCAACCAAGCCAGCAGCAATGATGCAGATGTCACCCAGACGATGGCCA 180
 QY 61 ThrTrpValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuLeaIaTrpLeu 80
 |||||
 Db 181 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGTAACTGGCTTATCTC 240
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
 |||||
 Db 241 GGACATTACCAACGCGCGCTGTGATAGCCGACGAGATGCTGGCGAAAAAGGCTGCCCC 300
 QY 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgIleGlyGlu 120
 |||||
 Db 301 CAGTCTGTCAGGTCCGCGATTATGCTGATGTCGACGAGGAGTACCCGTAACAGCGGAA 360
 QY 121 AlaPheAlaIleGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
 |||||
 Db 361 GCGTGGCCGCGCGGCTGGACCTGACCTGCAATACCGTACATACCCAGCGAGATACG 420
 QY 141 SerSerProAspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAla 160
 |||||
 Db 421 TCCAGTCCCGATCCGTTATTTATCTCTAATAAACCTGGGTTGGCCAACTGATACGCG 480
 QY 161 AsnValThrAspAlaIleLeuSerArgAlaGlyGlySerIleAlaAspPheThrGlyHis 180
 |||||
 Db 481 AACGTACTGACCCGCTCCTCAGCAGGCGAGGAGGCTCAATTCCTGACTTACCGGCGAT 540
 QY 181 ArgGlnThrAlaPheArgGlnLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
 |||||
 Db 541 CGGCAAAAGCGCGTTTGGCGACACTGGACGAGGAGGCTTAAATTTCCGCAATCAACCTGTGC 600
 QY 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
 |||||
 Db 601 CTTAAACGTGAGAAACAGGACGAACCTGTTCATTAAACGAGCATTTACATCGGAACTC 660
 QY 221 LysValSerIleAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
 |||||
 Db 661 AAGGTAGGCGCCACATGCTCATTAACCGGTGCGTAACCTCCATCAATGCTGACG 720
 QY 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
 |||||
 Db 721 GAGATTTTCTCTCGACACACACAGGAGATCCCGGAGCGGGGGGAAAGATCACC 780
 QY 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTrpLeuGln 280
 |||||
 Db 781 GATTACACACAGTGAACACCTTGAAGTTGTCATTAACGCGCAATTTATTTGCTACAA 840
 QY 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
 |||||
 Db 841 CGCACGCCAGAGGTTCCCGCACGCCGCCCGCTTATGGAATTTGATCAATGAGGACGC 900
 QY 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
 |||||
 Db 901 TTGACGCCCATCCACCGCAAAAACAGCGGTATGGTGTGACATTTACCCACTTGTAGTACTG 960
 QY 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyValAlaLeuGluLeuAsnTrp 340
 |||||
 Db 961 TTTATATGCCGACACATTAATCTGGCAATCTCGCGCGGCGCACTGGAGCTCAACTGG 1020
 QY 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyValLeuValPheGluAlaGlyTrp 360
 |||||
 Db 1021 ACCGCTTCCGCGTACGCGGATTAACACGCGCGCGCTGAACTGCTGTTTGAACCTGG 1080
 QY 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnTrpLeuGln 380
 |||||
 Db 1081 CCGTGGCTTAAGCATTAACAGCAGTGTATGAGTTTCCGTGCTTCCAGATTTTACAG 1140
 QY 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValValLysLeuThr 400
 |||||
 Db 1141 CAGATCGGTATTAACGCGCGTGTATTAATTAATACCGCGCGCGAGAGGTGAACACTGACC 1200

QY 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
 |||||
 Db 1201 CTGCGAGGATGTGATAGCGCAAAATGGCGACGCGATGTGCTTGGCAGGTTTACGCCAA 1260
 QY 421 IleValAsnGluAlaArgIlePheProAlaCysSerLeuAspSerHisHisHis 440
 |||||
 Db 1261 ATCGTAATGAACACGACCATACCGCGGTGCGATTGAGATCTCATCACTACCATCAC 1320
 RESULT 4
 US-09-866-379-7
 ; Sequence 7, Application US/09866379
 ; Patent No. US20020136754A1
 ; GENERAL INFORMATION:
 ; APPLICANT: DIVERSA CORPORATION
 ; APPLICANT: SHORT, Jay
 ; APPLICANT: KREY, Keith
 ; APPLICANT: GRAY, Kevin
 ; APPLICANT: BARTON, Nelson
 ; APPLICANT: GARRETT, James
 ; APPLICANT: O'DONOGHUE, Eileen
 ; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHRYASES AND USES THEREOF
 ; FILE REFERENCE: DIVER1370-7
 ; CURRENT APPLICATION NUMBER: US/09/866,379
 ; PRIOR FILING DATE: 2001-05-24
 ; PRIOR APPLICATION NUMBER: US 09/580,515
 ; PRIOR FILING DATE: 2000-05-25
 ; PRIOR APPLICATION NUMBER: US 09/318,528
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: US 09/291,931
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: US 09/259,214
 ; PRIOR FILING DATE: 1999-03-01
 ; PRIOR APPLICATION NUMBER: US 08/910,798
 ; PRIOR FILING DATE: 1997-08-13
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 7
 ; LENGTH: 1901
 ; TYPE: DNA
 ; ORGANISM: Escherichia coli
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)..(1901)
 ; OTHER INFORMATION: n is any nucleotide
 ; US-09-866-379-7
 Alignment Scores:
 Pred. No.: 6,62e-256 Length: 1901
 Score: 2235.00 Matches: 430
 Percent Similarity: 99.54% Conservative: 0
 Best Local Similarity: 99.54% Mismatches: 2
 Query Match: 97.09% Indels: 0
 DB: 10 Gaps: 0
 US-09-777-566a-2 (1-440) x US-09-866-379-7 (1-1901)
 QY 1 MetLysAlaIleLeuIlePheProLeuSerLeuLeuIleProLeuThrProGlnInsAla 20
 |||||
 Db 188 ATGAAGACGATCTTAATCCATTTTATCTCTGTGATTCGTTAAACCCGCAATGTGCA 247
 QY 21 PheAlaGlnSerLurProGluLeuLysLeuGlnSerValValIleValSerArgHisGly 40
 |||||
 Db 248 TTGCGTCAAGTAGCGGAGCTGAAGCTGGAAAGTGTGTGATGTCTCAAGTGTCAAGT 307
 QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
 |||||
 Db 308 GTGCGTGTCTCAACCAAGCCAGCAGCACTGATGTCAGAGATGTCACCCAGACGATGGCCA 367
 QY 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuLeuLeaIaTrpLeu 80
 |||||
 Db 368 ACCTGGCCGGTAAACCTGGGTGGCTGACACCGCGNGGTGGTGAAGTAACTGCTATCTC 427
 QY 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100


```

|||||
428 GGACATTACCAACGCGACCTCTGTGTACCGACGATTGTCGGCAAAAAGGGCGCCG 487
QY GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgIsthrGlyGlu 120
Db CAGTCTGGTCAGGTGGCATTTATGCTATGTCGCGAGCGACCGCATTAACAGCGCA 547
QY AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValIsthrGlnAlaAspThr 140
Db GCCTTCGGCGCGGGCTGGACACGTGCTGCAATACCGTACATACCCAGCAGCATACG 607
QY SerSerProAspProLeuPheAsnProLeuIsthrGlyValCysGlnLeuAspAsnAla 160
Db TCCAGTCCCGATCCGTTATTTAACTCTAAACACTGCGTTGGCACTGATTAACGCG 667
QY AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrIsthr 180
Db AACGTGACTGACGGCATCTCAGCAGGCGAGAGGGTCAATTGCTGACTTTACCGGGCAT 727
QY ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuGly 200
Db CGGCAAGCGCGTTTCGGAACTGGAAAGGCTTAAATTTTCCCAATCAAACTTGTGC 787
QY LeuIsthrArgGluIsthrGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db CTTAAACGTGAGAAACAGCAGAAAGCTGTTCAATTAACGACGCAATTCATCGAGACTC 847
QY LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db AAGGTGACGGCGGCAATGTCTCATTAACCGGTGGGTGAGCTGCTCATCATGCTGACG 907
QY GluIlePheLeuLeuGlnGlnAlaGlnIleMetProGluProGluIsthrGlyValIleThr 260
Db GAGATATTTCTCTGCAACAAAGCAGCAGGGAATCCGCGAGCCGGGTGGGAGAGATCAC 967
QY AspSerHisGlnIsthrPasnThrLeuLeuSerLeuHisAsnAlaGlnPheThrLeuLeuGln 280
Db GATTTCACACACAGTGGAAACACCTTGCTAGTTGCTTAACGCGCATTTATTTGCTACAA 1027
QY ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAla 300
Db CGCAGCGCCAGAGTTGGCCGAGCGCCGACCCCGTATTAAGATTTATCAAGACAGCG 1087
QY LeuThrProHisProProGlnIsthrGlnAlaIsthrGlyValIsthrLeuProThrSerValLeu 320
Db TTGACGCGCCCATCCACCGCAAAAACAGCGCATATGTTGACATTAACCACTTCACTGCTG 1147
QY PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGluLeuAsnThr 340
Db TTATATCGCGGAGACAGATACTAATCTGSCAAATCTCGCGCGGCGACTGAGACTCAACTCG 1207
QY ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValIsthrGluArgThr 360
Db ACGCTTCCCGGTGACCGCGGATACACGCGCGAGGTGTGAACGTGTTTGAACGCTGCG 1267
QY ArgArgLeuSerAspAsnSerGlnIsthrIleGlnValSerLeuValPheGlnThrLeuGln 380
Db CGTGGCTTAACGATTAACAGCAGTGAATTCAGGTTTGGCTGCTTCCAGACTTACAG 1327
QY GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGlyGlyValIsthrLeuThr 400
Db CAGATGCGTGAATAAAGCCGCTGTCATTAAATACGCGCGGAGAGTGAACATGAC 1387
QY LeuAlaGlyCysGluGluIsthrArgAsnAlaGlnIsthrMetCysSerLeuAlaGlyPheThrGln 420
Db CTGGCAGGATGTGAAGCAGAAATGCGAGGCGCATGTTGCTGCGAGGTTTATACGCA 1447
QY IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db ATCGTGAATGAAGCAGCATACCGGCTGACGTTTG 1483

```

```

; Sequence 9, Application US/09866379
; Patent No. US20020136754A1
; GENERAL INFORMATION:
; APPLICANT: DIVERSA CORPORATION
; APPLICANT: SHORT, Jay
; APPLICANT: KRETZ, Keith
; APPLICANT: GRAY, Kevin
; APPLICANT: BARTON, Nelson
; APPLICANT: GARRETT, James
; APPLICANT: O'DONOGHUE, Eileen
; TITLE OF INVENTION: RECOMBINANT BACTERIAL PHOSPHATASES AND USES THEREOF
; FILE REFERENCE: DIVERI370-7
; CURRENT APPLICATION NUMBER: US/09/866,379
; PRIORITY FILING DATE: 2001-05-24
; PRIOR APPLICATION NUMBER: US 09/580,515
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 09/318,528
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: US 09/291,931
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: US 09/259,214
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: US 08/910,798
; PRIOR FILING DATE: 1997-08-13
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 1901
; TYPE: DNA
; ORGANISM: Escherichia coli
; FEATURE:
; NAME/KEY: misc. feature
; LOCATION: (1)..(1901)
; OTHER INFORMATION: n is any nucleotide
US-09-866-379-9
Alignment Scores:
Pred. No.: 7,8e-255 Length: 1901
Score: 2226.00 Matches: 429
Percent Similarity: 99.318 Conservative: 0
Best Local Similarity: 99.318 Mismatches: 3
Query Match: 96.70% Indels: 0
DB: 10 Gaps: 0
US-09-777-566a-2 (1-440) x US-09-866-379-9 (1-1901)
QY 1 MetLysAlaIleLeuIlePheProLeuSerLeuLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGAGCATCTTATCCCATTTTATCTTCTGATTCGCTTAACCCCGCAATCTGCA 247
QY 21 PheAlaGlnSerGluProGluLeuIsthrGlnIsthrGlyValIleValSerArgHisGly 40
Db 248 TTCGCTCAGAGTACCCGAGCTGGAAGCTGGAAGTGTGCTATTGTCACTGTCATGTGT 307
QY 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValIsthrProAspAlaThrPro 60
Db 308 GTGGCTGCTCCACCAAGCCAGCAGCACTGATGACAGAGTGTACCCAGCAGCATGGCCA 367
QY 61 ThrTrpProValLysLeuGlyIsthrProLeuThrProArgGlyGlyGluLeuIleAlaThrLeu 80
Db 368 ACCGTGCGGCTGAACACTGGGTGGCTGACACCGCGNGGTGAGACTAATCCCTATCTC 427
QY 81 GlyHisThrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaIsthrGlyCysPro 100
Db 428 GGACATTACCAACGCGACGCTGTGTGACCGAGATTCCTGGCAAAAAGGGCGCCG 487
QY 101 GlnserglyValAlaIleIleAlaAspValAspGluArgThrArgIsthrGlyGlu 120
Db 488 CAGTCTGGTCAGGTGGCATTTATGCTATGTCGCGAGCGACCGCATTAACAGCGCA 547
QY 121 AlaPheAlaIleAlaGlyLeuAlaProAspCysAlaIleThrValIsthrGlnAlaAspThr 140
Db 548 GCCTTCGGCGCGGGCTGGACACGTGCTGCAATACCGTACATACCCAGCAGCATACG 607

```



```

|||||
Db 962 GATTACACACAGTGAACCTTGCTAGTTGCATTAACGGCCAAATTTATTACTACAA 1021
Qy 261 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
Db 1022 CGCACGCCAGAGGTTGCCCGCAGTCGCCACCGCCGATTTGATTGATCGACGCC 1081
Qy 301 LeuThrProHisProProGluLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1082 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGTACATTAACCCACTTCAGTGCCTG 1141
Qy 321 PheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaIleuGluLeuAsnTrp 340
Db 1142 TTTATTGCCGCGACAGATCACTATCTGCCAAATCTCCGGCGGCACATGACGTCACATCG 1201
Qy 341 ThrLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPheGluArgTrp 360
Db 1202 ACGCTTCACAGTCACCGCGATMACACGCCCGCAGGTGTGAACCTGCTTTGAACGCTGG 1261
Qy 361 ArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGlnThrLeuGln 380
Db 1262 CGTGCGCTTAAGCGATTAACAGCCAGTGCATTCAGTTGCTGCTTCACAGACTTACAG 1321
Qy 381 GlnMetArgAspLysTrpThrProLeuSerLeuAsnThrProProGlyGlyValLysLeuThr 400
Db 1322 CAGATGCGTGATTAACAGCCCGCTATCATTAATACGCCCGCGAGAGGTGAACACTGACC 1381
Qy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1382 CTGGAGAGATGTGAAGACCGAAATCGCAGAGGCATGTCTCGTGGCCGCTTTACGCCAA 1441
Qy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1442 ATCGTAATGAAGCCCGCATACCGCGCTGCAGTTTG 1477

RESULT 7
US-09-866-379-6
: Sequence 6, Application us/09866379
: Patent No. US20020136754A1
: GENERAL INFORMATION:
: APPLICANT: DIVERSA CORPORATION
: APPLICANT: SHORT, Jay
: APPLICANT: KREYFZ, Keith
: APPLICANT: GRAY, Kevin
: APPLICANT: BARRON, Nelson
: APPLICANT: GARRETT, James
: APPLICANT: O'DONOHUE, Eileen
: TITLE OF INVENTION: RECOMBINANT BACTERIAL PHYTASES AND USES THEREOF
: FILE REFERENCE: DIVER1370-7
: CURRENT APPLICATION NUMBER: US/09/866,379
: PRIOR FILING DATE: 2001-05-24
: PRIOR APPLICATION NUMBER: US 09/580,515
: PRIOR FILING DATE: 2000-05-25
: PRIOR APPLICATION NUMBER: US 09/318,528
: PRIOR FILING DATE: 1999-05-25
: PRIOR APPLICATION NUMBER: US 09/291,931
: PRIOR FILING DATE: 1999-04-13
: PRIOR APPLICATION NUMBER: US 09/259,214
: PRIOR FILING DATE: 1999-03-01
: PRIOR APPLICATION NUMBER: US 08/910,798
: PRIOR FILING DATE: 1997-08-13
: NUMBER OF SEQ ID NOS: 10
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 6
: LENGTH: 1901
: TYPE: DNA
: ORGANISM: Escherichia coli
: FEATURE:
: NAME/KEY: misc.feature
: LOCATION: (1)..(1901)
: OTHER INFORMATION: n is any nucleotide
US-09-866-379-6

```

```

Alignment Scores:
Pred. No.: 3 07e-254 Length: 1901
Score: 2221.00 Matches: 428
Percent Similarity: 99.07% Conservative: 0
Best Local Similarity: 99.07% Mismatches: 4
Query Match: 96.48% Indels: 0
DB: 10 Gaps: 0

US-09-777-566a-2 (1-440) x US-09-866-379-6 (1-1901)
Qy 1 MetLysAlaIleLeuIleProPheLeuSerLeuIleProLeuThrProGlnSerAla 20
Db 188 ATGAAGAGCATCTTAATCCATTTTATCTCTGATTCGTTAACCCGCAATGTGCA 247
Qy 21 PheAlaGlnSerGluProGluLeuLysGluSerValValIleValSerArgHisGly 40
Db 248 TTCGCTACAGATGAGCCGAGGCTGAAGCTGGAAAGTGGTGTGATTCAGTGTCAATG 307
Qy 41 ValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTrpPro 60
Db 308 GTGCGTCTCCAAACCAAGGCCACGCACTGATGACAGATGTCAACCCCAAGCATGGCCA 367
Qy 61 ThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTyrLeu 80
Db 368 ACCTGCGCGTAAACCTGGGTGGCTGACACCGCGNGTGTGAGCTTAATCGCCTATCTC 427
Qy 81 GlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysPro 100
Db 428 GGACATTAACCTGGGCCAGCGCTGTGTAGCGGAGTTCGTGGCGAAAGGCGTCCCG 487
Qy 101 GlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysTrpGlyGlu 120
Db 488 CAGTCTGTCAGAGTCCCGATTAATCTGATGTCGACGACGTCATCCCTTAACAGCGCGAA 547
Qy 121 AlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThr 140
Db 548 GCCTTGGCGCGCGGTGGCACCTGACCTGCAATTAACGGTACATACCCAGGACGATAGC 607
Qy 141 SerSerProAspProLeuPheAsnProLeuLysTrpGlyValCysGlnLeuAspAla 160
Db 608 TCCAGTCCCGATCCGTTATTTAACTCTTAACAAACGGGGTGGCCAACTGGATTAACGG 667
Qy 161 AsnValThrAspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHis 180
Db 668 AACGTACTGACGCCGATCTCTACAGAGGACGAGAGGGTCAATTCCTACTTTACCGGCGAT 727
Qy 181 ArgGlnThrAlaPheArgGluLeuGluArgValLeuAsnPheProGlnSerAsnLeuCys 200
Db 728 CGGCNAACGCGGTTTGGCGAAGTGGAAACGGGCTTAATTTCCGCAATCAAACTGTGCG 787
Qy 201 LeuLysArgGluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeu 220
Db 788 CTTAAACGTGAAGAACAGACGAAAGCTGTTCATTAAACGACGATTAACCATCGAATC 847
Qy 221 LysValSerAlaAspAsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThr 240
Db 848 AAGGTAGGCGCCGACAAATGTCTCATTAACCGGTCGGTAAAGCTCCCATCAATAGCTGACG 907
Qy 241 GluIlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTrpGlyArgIleThr 260
Db 908 GAGATATTCTCTCTGCAACAGACAGGAAATCCGAGCGGGGTGGGGAAGATCACCC 967
Qy 261 AspSerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnIleThrIleLeuGln 280
Db 968 GATTACACACAGAGAACACTTGTGAATTTGGCATTAACGCGAAATTTATTTGCTACAA 1027
Qy 281 ArgThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaIa 300
Db 1028 CGCACGCCAGAGGTTGCCCGCAGTCGCCACCGCCGATTTGATTGATCGACAGACGCC 1087
Qy 301 LeuThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSerValLeu 320
Db 1088 TTGAGCGCCCATCCACCGCAAAAACAGCGGTATGGTGTACATTAACCCACTTCAGTGCCTG 1147

```

OY	331	PheIlealaglYHhSAspThrAsnleuAlaasnleuglglYAlaLeuGluleuAsnTrp	340
Db	1148	TTTTATCCGCCGACACGACTACTAACTGTGCAAATCTCGCGCGCCACTGACCTCAACTGG	1207
OY	341	ThrlleuproglYglInProAspAsnThrProProglYglYgluleuValPhegluaRgTrp	360
Db	1208	ACGCTTCCCGGTCAAGCCGATTAACAGCGCCAGGTGGTGAACGTGTTTGAACGCGTGG	1267
OY	361	ArgArgleuSerAspAsnSerGlNTripIleGlnValSerleuValPheglInThrleuGln	380
Dc	1268	CGTCGGCTAACGATACAGCCAGCTGATTTAGCTTTGCTGCTGCTTCCACACTTTACAG	1327
OY	381	GlnMetArgAspLysTrhProleuSerleuAsnThrProProglYglYgluValLysleuThr	400
Db	1328	CAGATGCGTGATAAAGCCGCGCTGTCAATTAATACGCCGCCGGAGAGCTAAACTGACC	1387
OY	401	LeuAlaGlYcYsGlulGluaArgAsnAlaGlnGlyMetCysSerleuAlaGlyPheThrGln	420
Db	1388	CTGCAGAGATGTGAAGAGCCAAATGCGCAGCGCATGTGTCGTTGCGCAGGTTTACGCAT	1447
OY	421	IleValasnGluAlaArgIleProAlaCysSerleu	432
Db	1448	ATCGTGATGAAGCAGCATACCGGCGCGTGGCAGTTGG	1483

```
Db 1268 CGTGGCTAAGCATAAACAGCCAGTGCAGTTGCGTCTTCAGACTTTACAG 1327
Oy 381 GlnMetArgAspLysThrProLeuSerLeuAsnThrProProGluValLysLeuThr 400
Db 1328 CAGATGCGTGTATAAACCGCCCTGTCATTAAATAGCCGCCCGGAGAGTGAACGTGACC 1387
Oy 401 LeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGln 420
Db 1388 CTGGCAGATGTGAAGACGAAATCGCAGGCGCATGTGTTCTTGGCAGGTTTACGCA 1447
Oy 421 IleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1448 ATCGTGAATGAAGCAGCATACCGCGCGTCAAGTTTG 1483

RESULT 9
US-10-021-723A-11
; Sequence 11, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Mathur, Jay
; APPLICANT: Richardson, Toby
; APPLICANT: Richardson, Toby
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; CURRENT FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; PRIOR FILING DATE: 2000-12-12
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1326
; TYPE: DNA
; ORGANISM: Yersinia pestis
US-10-021-723A-11

Alignment Scores:
Pred. No.: 1,98e-104 Length: 1326
Score: 960.50 Matches: 202
Percent Similarity: 61.24% Conservative: 65
Best Local Similarity: 46.33% Mismatches: 154
Query Match: 41.72% Indels: 15
DB: 9 Gaps: 7

US-09-777-566A-2 (1-440) x US-10-021-723A-11 (1-1326)
Oy 5 LeuIleProPheLeuSerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSer 24
Db 37 TTAGTGTATATGCTCTCCGGCTGCGCTGCTATCACCGCGCCCTGTGGCC-----GCC 87
Oy 25 GluProGlu---LeuLysLeuGlnSerValIleValSerArgHisGlyValArgala 43
Db 88 GAACCATGCGGGGTACACCTTAGAACGTAGATTATTTAGAGCGCCAGTGGTGGTAC 147
Oy 44 ProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThrPro 63
Db 148 CCGACTAAGCAGACCCAGCTGTGTAATGATGAACACCTGATTAAGTGGCTCAGTGGCG 207
Oy 64 ValLysLeuGlyThrLeuThrProArgGlyGluLeuIleAlaIleAlaTyrLeuGlyHisTyr 83
Db 208 GTTAAAGCGGGCTATTACTCTCTGCGCGCCGAACTGGGTACCCCTGATGGCGGGTTC 267
Oy 84 GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProGlnSerGly 103
Db 268 TATGCGCATATATTCGCGAGTTTGGCTTTTGGCC---CGCGGGTGGCCCGCAGAGGGC 324
Oy 104 GlnValAlaIleIleAlaAspValAlaAspGluArgThrArgLysThrGlyGluAlaPheAla 123
Db 325 GGTGTATATGACAGACGAGATATGACACAGCGACCTGTTTAAACGGGTACAGCTTTTCTG 384
```

```
Oy 124 AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerPro 143
Db 385 GATGGTGTGGCCCGGGTGGCGGCTGCTGACTGTGCCAATGCAGCCGCTTAAAGAAACC 444
Oy 144 AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr 163
Db 445 GATCCTCTGTTTCATCCCGCTGAAACCGCGCTGTGAACCTGAGACACCCCAACCCGAT 504
Oy 164 AspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGlnThr 183
Db 505 AAGCAATTTGAGAAAGCCCTGGCGCGCCGTTAGACACCGTAAACCGCATATGCCA 564
Oy 184 AlaPheArgGluLeuGlnValLeuAsnPheProGlnSerAsnLeuCysLeuArg 203
Db 565 CCGTTTGGCAATATGGCGCATGTCTGAACTTCGCTGCAGAGTCCGACTGCAGACAGC 624
Oy 204 GlnLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
Db 625 CAGCAGACAGAGGAAAACTTGACTTCGCACACTTTCGCGCCACAGAAAGTAAATGTAAC 684
Oy 224 AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlu 241
Db 685 AAGGAAGCGCAGAAAGTTACCTGTACAGCCCGCCGCTGTCTTACACAGCTTGGCGCA 744
Oy 242 IlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrGlyArgIleThrAsp 261
Db 745 ATCTTCTTCTGTACAGACCGCAGCGATGATCCCGAAGTACGCTGCAGCGTTTGAAGGC 804
Oy 262 SerHisGlnTrpAsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArg 281
Db 805 GCTGAGAACCTGGGTGCTCTCTGAGCCCTGCACAAATGCACAGTCAACTGATGGCTA 864
Oy 282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeu 301
Db 865 ACGCCATACATTGCACGCCCAAGCAGCGCTTTTACACCAATGATACCGACATCG 924
Oy 302 ThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer----- 318
Db 925 ACC-----CTGCAACTGTGACCGCCCGGCGGCAAAACTGCGCATCTCGGCTCAGAAC 975
Oy 319 ---ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyGlyAlaLeuGln 337
Db 976 CGTGTATATCTCTGCTGGTGGCCAGCAGCAAAATATGCTAACATCGCCGCTATGCTGGGC 1035
Oy 338 LeuAsnTrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPhe 357
Db 1036 GCAGATGGCAGTTACCGGAAACCGCATACCCCGCCGCGCGCTGTGCTTT 1095
Oy 358 GluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGln 377
Db 1096 GAGCTGTGGCAAAATCCGCAATATCATCAAGCTTATGTGGCCGTTAAAGATGTTCTATCAG 1155
Oy 378 ThrLeuGlnGlnMetArgAspLysThrProLeuSerLeu---AsnThrProProGlyGln 396
Db 1156 ACCATGAGTCAATTTGGGTAAACCCGAGAGCTGATTTAAAGAACATCCCGCGGCATC 1215
Oy 397 ValLysLeuThrLeuAlaGlyGluGlnArgAsnAlaGlnGlyMetCysSerLeuAla 416
Db 1216 ATCAGTGTGCTGTGGCGCGCGGCGAATAATGTTGAGCATTAATGTGCAAACTGTAT 1275
Oy 417 GlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 432
Db 1276 ACTTTTCAAAAAAAGTACGCAAGTCAATTGAACCTGCTGTCAATTT 1323

RESULT 10
US-10-021-723A-3
; Sequence 3, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Mathur, Jay
; APPLICANT: Short, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Richardson, Dan
```

```

; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 1325
; TYPE: DNA
; ORGANISM: Versinia pestis
US-10-021-723A-3

```

```

Alignment Scores:
Pred. No.:      8,24e-102      length:      1325
Score:          938.50         Matches:      202
Percent Similarity: 61.248     Conservative: 65
Best Local Similarity: 46.338   Mismatches:  154
Query Match:    40.77%         Indels:      16
DB:              9            Gaps:        7

```

US-09-777-566a-2 (1-440) x US-10-021-723A-3 (1-1325)

```

QY      5 LeuileProphleuSerleuLeuileProleuThrProglInSerAlaPheAlaGlnSer 24
DB      37 CTGGTATTGATGCTAGACGGATGCGTGTATACGCGCGGTACCC-----GCC 87
QY      25 GluProGlu---LeuylSerGluSerValIleValSerArgHisGlyValArgAla 43
DB      88 GACCCATCGCGGCTATCTTACGACGTGGGTATTTTGGATGCGCATGGTGGTCTCG 147
QY      44 ProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaThrProThrPro 63
DB      148 CCAGCAACAAACA-ACGAGAGTTATGATGATGATGATGATGATGATGATGATGATGATG 206
QY      64 ValLysLeuGlyThrLeuThrProArgGlyGlyGlyLeuLeuAlaTyrLeuGlyHisTyr 83
DB      207 GTAAAGGCGGGGATTAAAGCGCCAGCGGTGGCGAGTGGTCAATGATGGGGGCTTT 266
QY      84 GlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCysProGlnSerGly 103
DB      267 TATGTCGATTACTTTCGACGCTTGGTTGTATGCG---GCGGAGATGCGCGCAGAGGG 323
QY      104 GlnValAlaIleIleLeuAlaAspValAspGluArgThrArgLysThrGlyGluAlaPheAla 123
DB      324 GGGGTATATGCACAGCAGCATATGCATCAACGATACCGGTTACGCGACAGCATTTCTT 383
QY      124 AlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAspThrSerSerPro 143
DB      384 GATGGTGGTGGTCCGCGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 443
QY      144 AspProLeuPheAsnProLeuLysThrGlyValCysGlnLeuAspAsnAlaAsnValThr 163
DB      444 GATCCCTTTCATCCGTCAGACTGGCGTGTGTGTAAGTTACACACGACAAACAGAT 503
QY      164 AspAlaIleLeuSerArgAlaGlySerIleAlaAspPheThrGlyHisArgGlnThr 183
DB      504 AAAGCATGTAAGAACGATTGGCGCGCGGTATGATAGCGTACGACAGCGCTACGTA 563
QY      184 AlaPheArgGluLeuGlnArgValIleAsnPheProGlnSerAsnLeuGlySerValArg 203
DB      564 CCTTTTGGCCAGTGGGAGCGTGAATTTTGGCGTTCCTTATTCGCAATCTTTG 623
QY      204 GluLysGlnAspGluSerCysSerLeuThrGlnAlaLeuProSerGluLeuLysValSer 223
DB      624 CAACACCAAGGAAACGCTGTATTTGCCACCTTTGGCGCCACATGATTAAGTTAAT 683
QY      224 AlaAsp-----AsnValSerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGlu 241
DB      684 AAGAAGGAGCAAAAGTACCTCACTGGCGCCACTGGCTTATCATCGACATTTGGGTGA 743

```

```

QY      242 IlePheLeuLeuGlnGlnAlaGlnGlyMetProGluProGlyTyrPglArgIleThrAsp 261
DB      744 ATCTGTATTATTAACAACGACCAAGCATGCCAGATGCCAGATGCCAGATGCCAGATGCC 803
QY      262 SerHisGlnThrAsnThrLeuSerLeuLeuSerLeuHisAsnAlaGlnPheThrLeuGlnArg 281
DB      804 GCGGAGAAATGGGATCATTTATTCATATACATACGCGCAATTAATTAATTAAGGCAAA 863
QY      282 ThrProGluValAlaArgSerArgAlaThrProLeuLeuAspPheLeuAlaLeuAlaLeu 301
DB      864 ACACCGTATATGCCCCGCTCATTAAGGAGCCCATTTTACAGACGATATGATGCGCTTTA 923
QY      302 ThrProHisProProGlnLysGlnAlaTyrGlyValThrLeuProThrSer----- 318
DB      924 ACC-----CTTCAACGATGATGCTCAGGGGCAAAAGCTACCTTACGCCCAAAAC 974
QY      319 ---ValLeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGlu 337
DB      975 CGGGTCTTGTCTCCGTCGGTGGCATGATACCAATATTTGCCATATTCGGGTATGTCGGA 1034
QY      338 LeuAsnThrProLeuProGlyGlnProAspAsnThrProProGlyGlyGluLeuValPhe 357
DB      1035 GCGGACTGGCAGCTACCCGACGACCACTGATATACACACGACGATGGGGGATTTGTTT 1094
QY      358 GluArgTrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerLeuValPheGln 377
DB      1095 GAACATATGGCAAAATCCGATATACACACGACGCTACGCTGGAAGATGTTCCACAA 1154
QY      378 ThrLeuGlnGlnIleArgAspLysThrProLeuSerLeu---AsnThrProProGlyGlu 396
DB      1155 ACAGATGATCAGTTCGCTATATGCGGAAATTTGATCTGAAATTAATACCGACGGGTATT 1214
QY      397 ValLysLeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnLysMetCysSerLeuAla 416
DB      1215 ATTCCGTCGACACTGCTGCTGTGTAATTAACGTCAGCATTAAGCTTTCGAGCTTGAT 1274
QY      417 GlyPheThrGlnIleValAsnGluAlaArgIleProAlaCysSerLeu 432
DB      1275 ACATTCCAAAAGAAAGTGGCTAAGTAAATTAATTAAGCTTCGCCACATC 1322

```

RESULT 11

```

US-10-021-723A-1
; Sequence 1, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1325
; TYPE: DNA
; ORGANISM: Versinia pestis
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 470, 487, 491, 492, 522
; OTHER INFORMATION: n = A,T,C or G
US-10-021-723A-1

```

```

Alignment Scores:
Pred. No.:      1,98e-99      Length:      1325
Score:          918.50         Matches:      199
Percent Similarity: 60.32%     Conservative: 64
Best Local Similarity: 45.64%   Mismatches:  158

```



```

Db      304 GATCGTGGCGCGCGCTGCTTCGACGATGGGGGGGTACTACGGTATGCTTCGTGCC 363
Oy      91 AspGlyLeuLeuAlaLysLysGlyCysProGlnSerGlyGlnValAlaIleIleAlaAsp 110
Db      364 GGGGGTGTGCTCCGCGCCACGCGGGTCCCTGGCAGCCGGAACCTTTACGATGGCGAGAC 423
Oy      111 ValAspGluArgThrArgLysThrGlyGlnAlaIleAlaGlyLeuAlaProAspCys 130
Db      424 GTTGGACACGCGGACGCTGTGACGGCGGACGCCCTCTCCGTGGCATGGCCGACGCTGC 483
Oy      131 AlaIleThrValHisThrGlnAlaAspThrSerSerProAspProIleuAsnProIleu 150
Db      484 GGTATCCACAGCGATCATCGCGGGCGCTCCAGCAGAACGATCCGATCTTCACAGCGATG 543
Oy      151 LysThrGlyValLysGlnIleuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAla 170
Db      544 GAATGGGGTGCATGGCCAGTCGACCCCGTCACAGCGCAACGGGACATCGAACCGCATGCC 603
Oy      171 Gly---GlySerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGln 189
Db      604 GCGGAGGGGGCGGTGGCGACACTGGGAGGCGCTACGACAGCAAGCCTGACCAAGATGAGC 663
Oy      190 ArgValIleuAsnPheProGlnSerAsnIleuCysIleuLysArgGluLysLinspGlnSer 209
Db      664 GAGGTGCTCGACTACGCCCATAGCCCGCATTTGCGGAGGCGATGGCGCCAA----- 714
Oy      210 CysSerLeuThrGlnAlaLeuProSerGlyLeuLysValSerAlaAsp-----AsnVal 227
Db      715 TCGACATATGGCGCCAA---CCCAATGCTGTGCAGATACAGACAGATGGCTTCATGCC 771
Oy      228 SerLeuThrGlyAlaValSerIleuAlaSerMetLeuThrGlnIlePheLeuGlnGln 247
Db      772 GCGTTGAAGGGCCGACAGTGGCAGTCTTCGACCGCTTCGACGAGTCTTCGCTCGCAACAT 831
Oy      248 AlaGlnGlyMetProGlnPro-----GlyTrpGlyArgIleThrAspSerHisGlnTrp 265
Db      832 GGGCAGAGGCGCTCCACAGAGCAGGTGGATGGGCGGTATCCAGATGCGACAGCATGG 891
Oy      266 AsnThrLeuLeuSerIleuHisAsnAlaGlnPheTrpLeuLeuGlnArgThrProGluVal 285
Db      892 ACGCTGCTGATGACAGCGCATGACGCGAGTTCGATGTCATGTCGAGCAAGCGCTTACATG 951
Oy      286 AlaArgSerArgAlaThrProIleuLeuAspIleMetAlaIleAlaLeuThrProHis--- 304
Db      932 GCTACTCGAAGGCGACCGCATGCTCGCTGGTGGTGGATGCGCTTGACGGCGCGCT 1011
Oy      305 -----ProProGlnLysGlnAlaTrpGlyValThrLeuPro-----ThrSerVal 319
Db      1012 GCGCGCCACAGCTCGGAGCTTCCCTCAAGGCCCGGACGCTGCCCAAGCGCAACGCTGTC 1071
Oy      320 LeuPheIleAlaGlyHisAspThrAsnLeuAlaAsnLeuGlyAlaLeuGluLeuAsn 339
Db      1072 TATGTGCTGACCGCGCATGACAGATCTTGCAACTTGGCGGGCTGCTGCACACTGGAC 1131
Oy      340 TrpThrLeuProGlyGlnProAspAsnThrProProGlyGlyLeuValPheGluArg 359
Db      1132 TGGACCTGCGCGAGCAACCGGACAGCGCGCGGGGTGTCATGCTGCTGCTTG 1191
Oy      360 TrpArgArgLeuSerAspAsnSerGlnTrpIleGlnValSerIleuValPheGlnThrIleu 379
Db      1192 TGGCGGGAGCGCGGACGACGACGCTTCCCTGCGGTGGAGATGGTCTATGATGCTGATG 1251
Oy      380 GlnGlnMetArgAspLysThrProLeuSerIleuAsnThrProProGlyValValLysLeu 399
Db      1252 GATCAGCTTCGCGACATCAGCGCGCTCCCTGGGGGAGCGCGCCCATGGCTGATCTTG 1311
Oy      400 ThrIleuAlaGlyCysGlnGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThr 419
Db      1312 CCGTTGGCCCGGTGTGGCGCGCGCGACGACATGATGATGATGATGATGATGATGATGATG 1371
Oy      420 GlnIleValAsnGluAlaArgIleProAlaCys 430

```

```

Db      1372 CGCGGTGTGCGCGCGCATTTGTCCTCCCTCTCTGC 1404
RESULT 13
US-10-021-723A-7
; Sequence 7, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Short, Jay
; APPLICANT: Mathur, Eric
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phytases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIOR FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 7
; LENGTH: 1230
; TYPE: DNA
; ORGANISM: Rhizobium
; US-10-021-723A-7

Alignment Scores:
Pred. No.: 7, 27e-47
Score: 476.50
Percent Similarity: 45.548
Best Local Similarity: 31.358
Query Match: 20.708
Gaps: 15

US-09-777-566A-2 (1-440) x US-10-021-723A-7 (1-1230)
Oy      10 SerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerGluPro----- 26
Db      10 AGCCTGCTTTGCTGCTCCGCGCTGCGGGGATGCGCGCGCGCGCGCGCGCGCGCGAGC 69
Oy      27 -----GluLeuLysLeuGlnSerValValIleValSerArgHis 39
Db      70 GCATGCGCTCCGCGCGCGCGGGGGGCTGTAAGCTCGAAGAGCTGTAACTGATGCTGATGCC 129
Oy      40 GlyValAlaArgAlaProThrLysAlaThrGlnIleuMetGlnAspValThrProAspAlaTrp 59
Db      130 GCGGTGCGCGCGCAAGCGGGGGTGGTCCCGCGGTATTCGCGCGCAACATGG 189
Oy      60 ProThrTrpProValLysLeuGlyTrpLeuThrProArgGlyGlyLeuIleAlaTrp 79
Db      190 CCCGACTGCGCGGTGATTTGCGCTGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCTT 249
Oy      80 LeuGlyHisTrpGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
Db      250 CTCGGCGAAGAGCGCGCTCTATTTCCGCGGCTCGCGCGCTGTT---CCCGAGGATGC 306
Oy      100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspGluArgThrArgLysThrGly 119
Db      307 CCGCGCGCGCGACGATGCTCTCAAGCGAGCTACAGAGCGCGACGATGCGCACCGCG 366
Oy      120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db      367 CAGAGCTGGGCGCGGTTCATGCGCGCTCGCGCGCGCGCGCGCGCGCGCGCGATGTC 417
Oy      140 ThrSerSerPro-----AspProLeuPheAsnProLeuLysThrGlyValCysGlnIleu 157
Db      418 CCCGCGCGGTCCGAGAGATGACGATCTTTCATGGGCTGACAGCGCGCGCGCGCGCTTC 477
Oy      158 AspAsnAlaAsnValThrAspAlaIleLeuSerArgAla-----GlyGlySerIleAla 175
Db      478 GACGGCAAGCGGGCGGTGATGCGCGCGTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 537
Oy      176 AspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValIleuAsnPhePro 195

```



```

Db      538 GAAACCGACGTCATCCG---GCCGACTGACCTTCTCGCGAAGTGTGTAAT----- 588
Oy      196 GlnSerAsnLeuCysLeuLysArgLysGlnAspGlnSerCysSerLeuThrGln---- 214
Db      589 -----TGGCGCTGCGCCCGCTGC 606
Oy      215 -----AlaLeuProSerGluLeuLysValSerAla---AspAsnValSerLeuThr 230
Db      607 CCGCGATCGCGAGCGGCGGCTGTGTCGGCGGACCGCGACGATCGCCCGCAGATCGGA 666
Oy      231 GAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGlnAlaGlnGly 250
Db      667 GCGCGCTCGACGTGCGATCGACCGCCAGACGCGTGTCTGGAATATCTGGAAGGC 726
Oy      251 MetPro-----GluProGlyTyrProLysArgIleThrAspSerHisGlnTyrAsnThrLeu 268
Db      727 AAGCGGATGGCGAGTGGCTGGGGCGCGCTGAGCGCCGCGC---GAGATCGAGCAGATTG 783
Oy      269 LeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgTyrProGluValAlaArgSer 288
Db      784 CTGGCGCTTCATCCGCTCAATTCGCGTATTCGAAATCCGCCGCTATATCCGCCGACAC 843
Oy      289 ArgAlaThrProLeuLeuAspLeuIleMetAlaAlaLeuThrProHisProGlnLys 308
Db      844 GCCCGCGCGCGATCGTCCGGAATCGTCAAGCGCGCTC----- 882
Oy      309 GlnAlaTyrGlyValThrLeuProThrSerValLeuPheIleAlaGlyHisAspThrAsn 328
Db      883 -----GGCGACCGAAGCCCGCGCGCTGACCTTGTCTCGCGGCGACACACGAAAC 933
Oy      329 LeuAlaAsnLeuGlyAlaLeuGlnLeuLeuAsnTyrProThrLeuProGlyLysPro---Asp 347
Db      934 GTCGCGGACCTCGCGGCTTCTTCGACCTGCACCTGCGACGATCCGAGTTATATCCGCCGAC 993
Oy      348 AsnThrProProGlyGlyLeuValPheGlnArgTyrArgArgLeuSerAspAsnSer 367
Db      994 GAGTTCGCGCGCGACGCGACCTGGGGTTGAA-----CTGGTCAACAATGGC 1041
Oy      368 -----GlnTyrIleGlnValSerLeuValPheGlnThrLeuGlnGlnMetArgAsp 384
Db      1042 AAGCGGACCGCTATGCTCCGCGCTTCATTCGCGGCGAGCATGACGACGACGCTCGGCAAC 1101
Oy      385 LysThrProLeuSerLeuAsnThrProProGlyGluValLysLeuThrLeuAlaGlyCys 404
Db      1102 CTCGACGCGCTGAGGTGCGAGCATACGCTGTCCCGCTATCTTCCCATTCGCGGGTGC 1161
Oy      405 GluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGlyPheThrGlnIle 421
Db      1162 GGCATTCGCTCGACGAGCGACCGCTGCGCTGAGTGTGATTCGCCCGGCTC 1212

RESULT 14
US-10-021-723A-5
; Sequence 5, Application US/10021723A
; Publication No. US20030101476A1
; GENERAL INFORMATION:
; APPLICANT: Mathur, Jay
; APPLICANT: Short, Jay
; APPLICANT: Richardson, Toby
; APPLICANT: Robertson, Dan
; APPLICANT: Barton, Nelson
; TITLE OF INVENTION: Recombinant Phylases and Uses Thereof
; FILE REFERENCE: 112766.140 (DIV-016CIP)
; CURRENT APPLICATION NUMBER: US/10/021,723A
; PRIORITY FILING DATE: 2002-10-23
; PRIOR APPLICATION NUMBER: US 60/255,090
; NUMBER OF SEQ ID NOS: 74
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Artificial Sequence

```

```

; FEATURE:
; OTHER INFORMATION: 953-6 phylase sequence
US-10-021-723A-5

Alignment Scores:
Pred. No.: 2,27e-46 Length: 1266
Score: 472.50 Matches: 137
Percent Similarity: 44.55% Conservative: 59
Best Local Similarity: 31.14% Mismatches: 177
Query Match: 20,538 Indels: 67
DB: 9 Gaps: 15

US-09-777-566a-2 (1-440) x US-10-021-723A-5 (1-1266)

Oy      10 SerLeuLeuIleProLeuThrProGlnSerAlaPheAlaGlnSerLysPro----- 26
Db      46 ACCCTGCTTTTGTTCGCGCGCGCTGCTGGGGGATGCGCGCTCCCGCGCGCGCGAC 105
Oy      27 -----GluLeuLysLeuGlnSerValValIleValSerArgHis 39
Db      106 GCATCGCCCTGCGCGCGCGCGGGGTGCGTAAAGCTCGAAGAAAGTGTCTGATGCGCCAC 165
Oy      40 GlyValArgAlaProThrLysAlaThrGlnLeuMetGlnAspValThrProAspAlaTyr 59
Db      166 GCGGTTCGCGCGCGCAAGCGGGGTGTGCGCGCGGTATTCGCGCGCAACATCG 225
Oy      60 ProThrTyrProValLysLeuGlyTyrThrLeuThrProArgGlyGlyLeuLeuIleAlaTyr 79
Db      226 CCGGACTGGCGCGGTGATTTGCGCTGCTGACGCGCGACGCGCGGGGGGTCAAGCTG 285
Oy      80 LeuGlyHisTyrGlnArgGlnArgLeuValAlaAspGlyLeuLeuAlaLysGlyCys 99
Db      286 CTCGCGGAAGGAGCGCGCTCACTTGGCGGTGCGCGCTATTC---CGGACGGTTGT 342
Oy      100 ProGlnSerGlyGlnValAlaIleIleAlaAspValAspLysArgTyrArgLysThrGly 119
Db      343 CCGCGCGCGGACGATCTCTCAAGCAACTACAGAGCGACGATCGCGACGCGG 402
Oy      120 GluAlaPheAlaAlaGlyLeuAlaProAspCysAlaIleThrValHisThrGlnAlaAsp 139
Db      403 GAAACTGGCGCGGGGTTCATGCCGCTGCACG-----GCGGAT 444
Oy      140 ThrSerPro-----AspProLeuPheAsnProLeuLysThrGlyVal 154
Db      445 GTCGCGGACCGCGCGGTCCGAGCATGACGAGATCTTCATGCGGTCCACGCGGCGCC 504
Oy      155 CysGlnLeuAspAsnAlaAsnValThrAspAlaIleLeuSerArgAla-----GlyGly 172
Db      505 GCCTGCTTCGACGCGGAGCGGCGCATTCGATCGCGCTCGCCAGCGCGCGAGGGCGGG 564
Oy      173 SerIleAlaAspPheThrGlyHisArgGlnThrAlaPheArgGluLeuGluArgValLeu 192
Db      565 CTGACCGCGGAAACCGCACCGCATTCG---GGCGAATGACCTTGTCCGGAAGTGTG 621
Oy      193 AsnPheProGlnSerAsnLeuLysLeuLysArgLysGlnAspLysSerCysSerLeu 212
Db      622 AAT-----TGTGGCTG 633
Oy      213 ThrGln-----AlaLeuProSerGluLeuLysValSerAla---AspAsnVal 227
Db      634 CCGCGCTCGCGCTGATCGCGAGCCGACCTGCTGCGCGACCGCGACGATCGCCCG 693
Oy      228 SerLeuThrGlyAlaValSerLeuAlaSerMetLeuThrGluIlePheLeuLeuGlnGln 247
Db      694 GATCTGGAAGCGCGCTGCGATCGGATGACCGCGGCAACACGCTGCTGCGGAATAT 753
Oy      248 AlaGlnLysMetPro-----GluProGlyTyrGlyArgIleThrAspSerHisGlnTyr 265
Db      754 CTGGAAGCGAAGCGCATGCGCGAGTGGCTGGGGCGGTGAAGCGCGCC---GAGATC 810
Oy      266 AsnThrLeuLeuSerLeuHisAsnAlaGlnPheTyrLeuLeuGlnArgTyrProGluVal 285
Db      811 GAGCAGTGTGCTCGCTTCATTCGCTCAATTCGCGTATTCGAATCGCCCGGCTATATC 870

```



```

Db      1225  ||:::|||  ::|||  |||||  ::|
          ACTATGCAATGTAAGTCAGAAAGAACCTTTA---GTTAGAGCTTTGATTAAACGACAGA 1281
OY      398  LysLeuThrLeuAlaGlyCysGluGluArgAsnAlaGlnGlyMetCysSerLeuAlaGly 417
          ::|  |||  |||||::|  ::|  |||  |||  |||
Db      1282  GTTGTTCATTACATGGTTGATGTT---GACAAGTTAGGTAGATGTAAGTTAAACGAC 1338
OY      418  PheThrGlnIleValAsnGluAlaArg 426
          |||  ::|  ::::|  |||||
Db      1339  TTTGTCAGAGGTTTGTCTTGCGCCAGA 1365

```

Search completed: June 12, 2003, 13:21:12
 Job time : 275 secs

THIS PAGE BLANK (uspto)